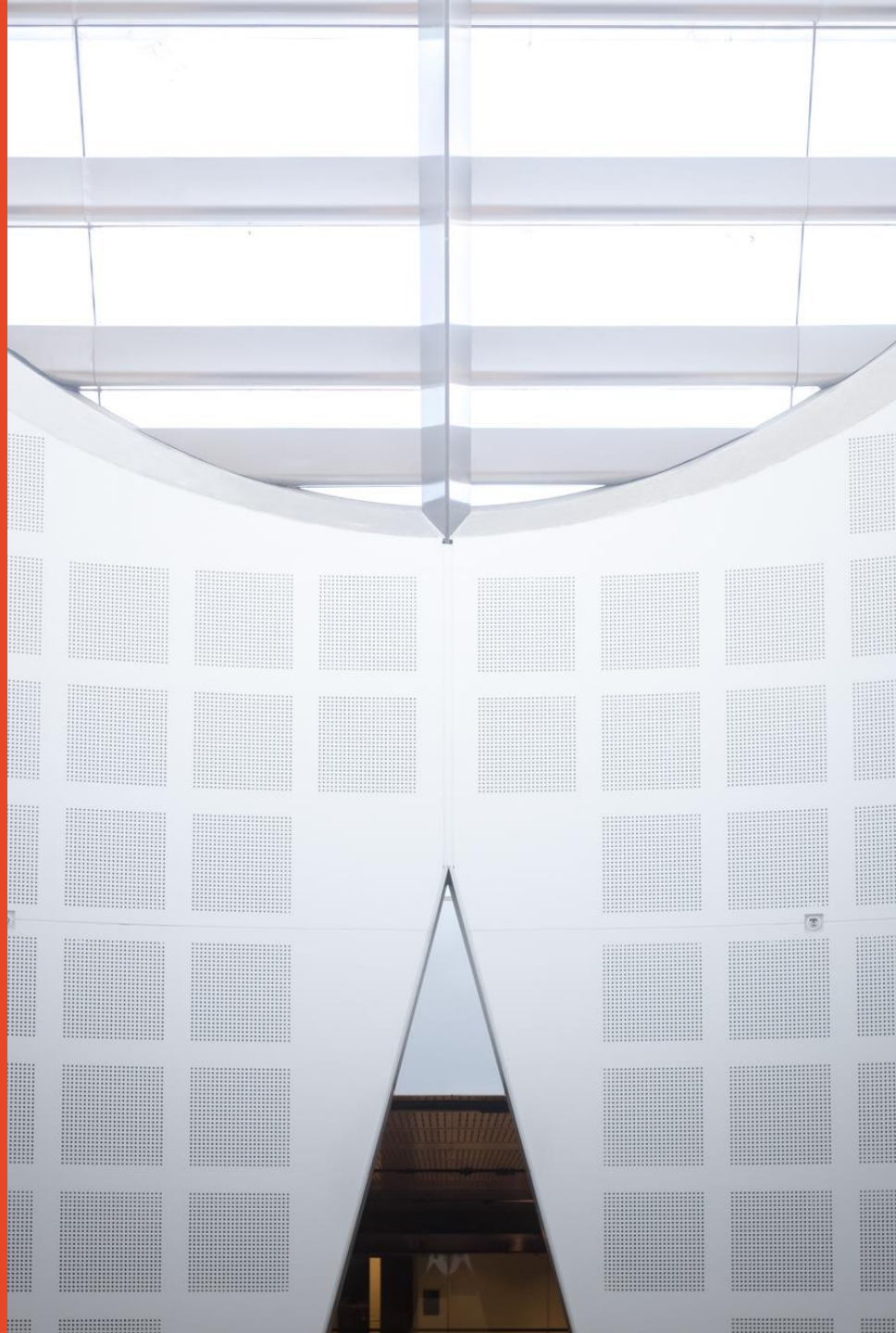
Meta-analysis
An Introduction

Presented by
Jim Matthews
Senior Consultant: Statistics
Sydney Informatics Hub
Core Research Facilities
The University of Sydney

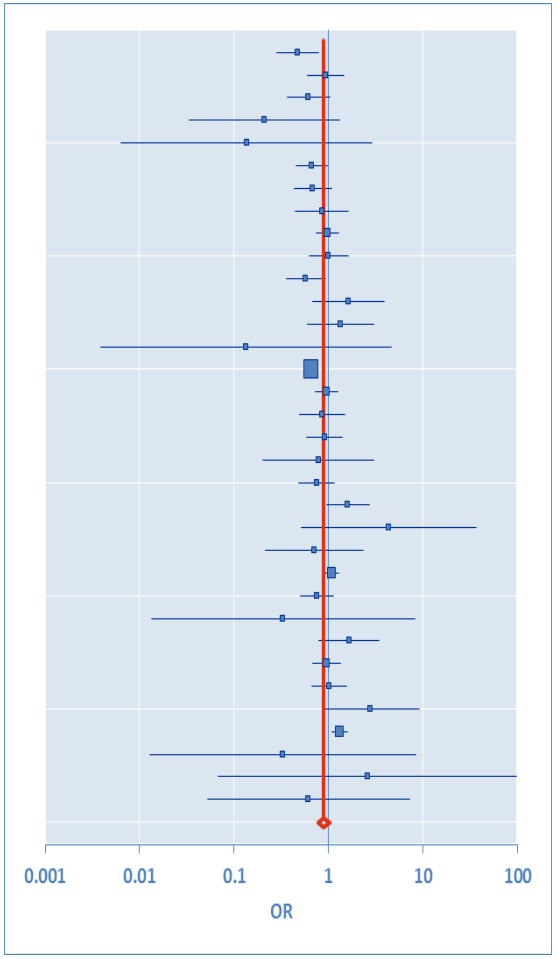
April 2021

THE UNIVERSITY OF SYDNEY



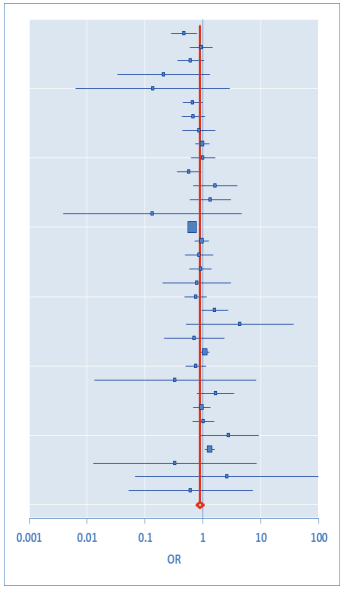
Outline

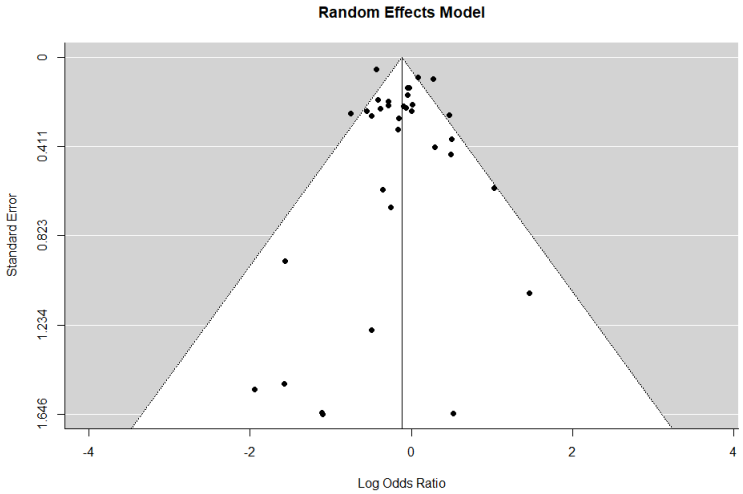
- The research context
- Systematic review – step by step overview
- Meta-analysis – concepts step by step
- Software options
- Worked example using metafor in R



Workshop Aims

- Understand the key concepts and tools in meta-analysis

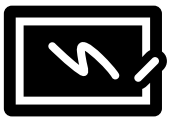




- Follow the steps to perform a meta-analysis
- Basic knowledge of performing these steps using a software package (metafor)

How to use this workshop

- These slides have a dual purpose:
  - To guide our interactive workshops
  - As self-contained reference material to be read after the workshop

- Some slides are for your reference, and not all of the material will be discussed in the workshop. Such slides are marked with this blackboard icon 

- Ask short questions or clarifications during the workshop. There will be breaks during the workshop for longer questions. You can email us about the material in these workshops at any time, or request a consultation for more in-depth discussion of the material as it relates to your specific project

Systematic Review - the research context

Narrative Review

This is generally what we see in a review article or in the first chapter of a research thesis. The aim is often to provide an overview rather than focussing on a single research question.

- Often subject to biases, in both selection and interpretation of studies, in order to fit a narrative.
- Methodology for surveying the topic and criteria for inclusion in the review are not defined.
- The studies might be treated differently according to the reviewer's assessment of their quality, however, this assessment will likely be subjective and unstated.
- Reasons for different findings are difficult to explore or resolve.

Systematic Review - the research context

Systematic Review

This aims to be a more objective review of a research topic.

- The research question is focussed and defined.
- Methodology for surveying the topic (e.g. databases, keywords) and criteria for inclusion in the review are defined in a protocol (eg PRISMA)
- Decisions on study inclusion, including assessing the quality of the study, are considered to be more objective.
- Information is extracted from the studies in a consistent manner.
- A **meta-analysis** (synthesis or combination of the data) is preceded by the systematic review. A meta-analysis doesn't always follow a systematic review.

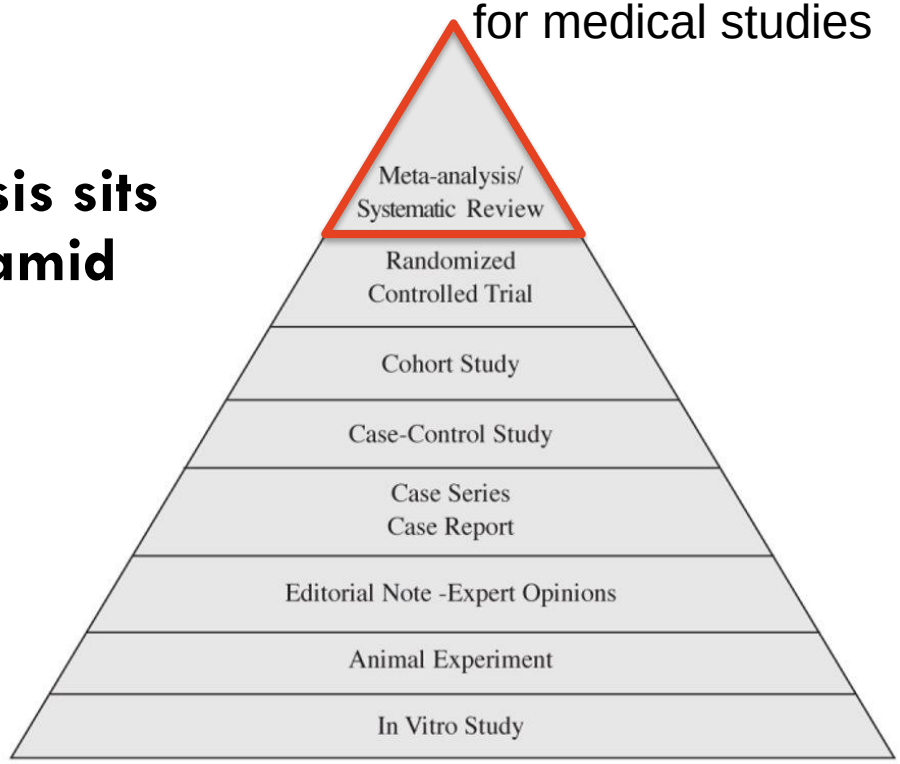
Systematic Review - the research context

Why choose to do Systematic Review and meta-analysis?

A well conducted meta-analysis sits at the top of the evidence pyramid

The results will be more generalisable than any single study

Evidence pyramid for medical studies

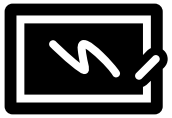


Systematic review – step by step overview

1. Formulate your research question
2. Develop eligibility criteria
3. Search the literature
4. Screen the literature, assessing the quality of studies
5. Extract the data
6. Synthesize and interpret the results
    - Meta-analysis
7. Publish and disseminate the results

The University of Sydney library guides
https://libguides.library.usyd.edu.au/c.php?g=516278&p=3529739

Literature Review – Library support



- The USyd library is the best place to get support on the wider questions of Literature Review and Systematic Review
- Resources available at the Library Subject Guides and Library Services
- Library has dedicated support staff – eligible staff and students can request a consultation with an Academic Liaison Librarian
- Workshops run through the year include "Getting started with Literature Review" – check the workshops calendar (filter category by "systematic review")
- Canvas modules available online

Meta Analysis – concepts and workflow step by step

Meta-Analysis workflow steps

1. Import the data
2. Calculate Effect Size
3. Meta-Analysis summary (Fixed-Effect and Random-Effects)
4. Explore Heterogeneity (Q statistic, I^2, tau)
5. Publication bias (Funnel plot)
6. Sub-group analysis and/or meta-regression (if warranted)

1. Import the data Roles of Variables

What type of data do we have?

A variable – any factor or measure that can vary.

Explanatory (Independent) Variables – For example,

- Something that is set as part of the experimental design (e.g. control group and treatment group).
- Factors in the experimental subjects that may be relevant to the study (e.g. age and sex).

Response (Dependent) Variables – these are measured or observed to determine whether they vary in response to the explanatory variables.

1. Import the data Types of variables

Numeric

Continuous
(ratio or interval)
- Weight/height
- Blood pressure
- Time
- Plasma glucose levels

Discrete (count)
- Number of myocardial infarctions
- Size of litter
- Number of visits to a cage corner.

Categorical

Ordinal
- Survey feedback
- Disease stages
- Pain descriptions

Nominal/binomial
- Eye colour
- Occupation
- Smoker/non-smoker
- Present/absent
- Success/fail

The type of variable will influence what type of effect size is appropriate.

1. Import the data Study questions

Because of the analysis methods, meta-analyses primarily address quite simple comparisons.

1. Continuous variable (e.g. weight loss) measured in two groups.
2. Binary events (e.g. myocardial infarction [heart attack] yes/no)
3. Correlation between two continuous variables (e.g. blood lead levels in children and IQ scores).

More complex analyses may look at the influence of a modifying factor(s) i.e. a categorical variable (subgroup effect) or continuous variable (meta-regression).

The types of variables and the study design determine the type of effect size that is synthesized in the meta-analysis.

2. Calculate Effect Size

Types - Difference in means

This is common when the explanatory variable is categorical and the response variable is numeric.

Example:

The study question is a comparison of the effectiveness of *low fat diets vs low carbohydrate diets for weight loss*

Explanatory variable is diet type – categorical (low fat or low carbohydrate). Response variable is weight loss – numeric (kg).

The effect size is the difference in mean weight loss between the two groups.

$$Effect\ size = \bar{x}_{LF} - \bar{x}_{LC}$$

2. Calculate Effect Size

Types – proportions and Risk Ratio

These are typically used where you have 2 independent groups with a binary outcome.

Example: The study question is a comparison of
survival rates in males and females suffering myocardial infarctions (MI)

Both variables are binary (male/female, death/survival).
p1= proportion of females dying from MI
p2= proportion of males dying from MI

$$Risk\ Difference\ Effect\ size\ = p1 - p2$$

Computations are carried out on the raw differences in proportions.

Another effect size (depending on the study design) is Risk Ratio

$$Risk\ Ratio\ Effect\ size = p1/p2$$

Computations are carried out on the natural log of the Risk Ratio.

2. Calculate Effect Size Types – Odds Ratio

The same outcome can also be expressed as an Odds Ratio.

	Death (event)	Survival (no event)	Proportion (Risk)	Odds
Female	a	b	p1=a/(a+b)	a/b
Male	c	d	p2=c/(c+d)	c/d

$$Odds\ Ratio\ Effect\ size = \frac{a/b}{c/d} = \frac{ad}{bc}$$

When the risk of the event (eg death) is low, then the Odds Ratio will be similar to the Risk Ratio.

Analysis of an OR effect size is often preferred to RR because of superior statistical properties.

Computations are performed on the natural log of the odds ratio.

2. Calculate Effect Size

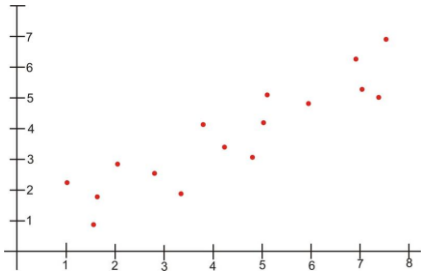
Types - Correlation

This is common when both variables are numeric.

Example:

Blood lead levels in children and IQ scores.

Both variables are numeric (IQ score and plasma lead concentration). We use the correlation coefficient.



$$Effect\ size \quad r = \frac{\sigma^2{}_{xy}}{\sigma_x \sigma_y}$$

2. Calculate Effect Size

Types - Correlation

Like with Odds Ratios, combining correlation coefficients needs to be done using the right scale. In this case we use the Fisher Z transformation

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

Study	(Pearson's) r	(Fisher's) z
A	0.20 →	0.2027
B	0.40 →	0.4236
Average (A:B)	0.3033 ←	0.3132

2. Calculate Effect Size

Combining Effect Sizes - considerations

Why not simply average the effect sizes from multiple studies?

There are a few things to consider

1. Standardisation
2. Weighting
3. Different types of effect sizes.

2. Calculate Effect Size

Standardised effect sizes

It may not be necessary to use standardised effect sizes.

Example 1: Selected studies examining effects of a supplement on fasting glucose levels may always measure the plasma glucose concentration.

Combine without standardisation

Example 2: Selected studies might examine ratings of anxiety among carers of dementia sufferers vs controls. Some studies used the Generalised Anxiety Disorder Questionnaire (GAD-7), while other studies used the Hamilton Anxiety Rating Scale (HAMA-A).

Standardise before combining

Standardised effect sizes are a way to transform effect sizes into a form where they can be compared and summarised.

2. Calculate Effect Size

Standardised effect sizes – difference in means

Many standardised effect sizes were defined by Jacob Cohen (1923-98). For example, Cohen's *d*

$$d = \frac{\overline{x_1} - \overline{x_2}}{s}$$

Where $s = \sqrt{\frac{[(n_1-1)s_1^2+(n_2-1)s_2^2]}{(n_1+n_2-2)}}$ (the pooled within group SD)

This index tends to be an overestimate when sample sizes are small (n<20) therefore the Hedges bias correction is often applied to calculate Hedges' *g*.

$$g = \left(1 - \frac{3}{4df-1}\right) d$$

$df=n_1+n_2-2$

2. Calculate Effect Size Weighting

Example:

The study question examines weight loss in a low carbohydrate diet (LC) vs a low fat (LF) diet.

Study 1 The difference in weight loss was 8 kg. The number of participants was 8 in each dietary group.

Study 2 The difference in weight loss was 0 kg. The number of participants was 64 in each dietary group.

Should each study contribute equally to the calculation of the overall effect size? Is the overall effect size a 4 kg weight loss?

2. Calculate Effect Size Weighting

- We need to consider the precision of the estimates for each study
- Studies that are more precise should have more weighting in the summary estimate
- Information about study precision is found in the Standard Error

2. Calculate Effect Size What is a standard error?

A measure of the precision of an estimate for a parameter.

Most commonly we think of the **standard error of the mean.**

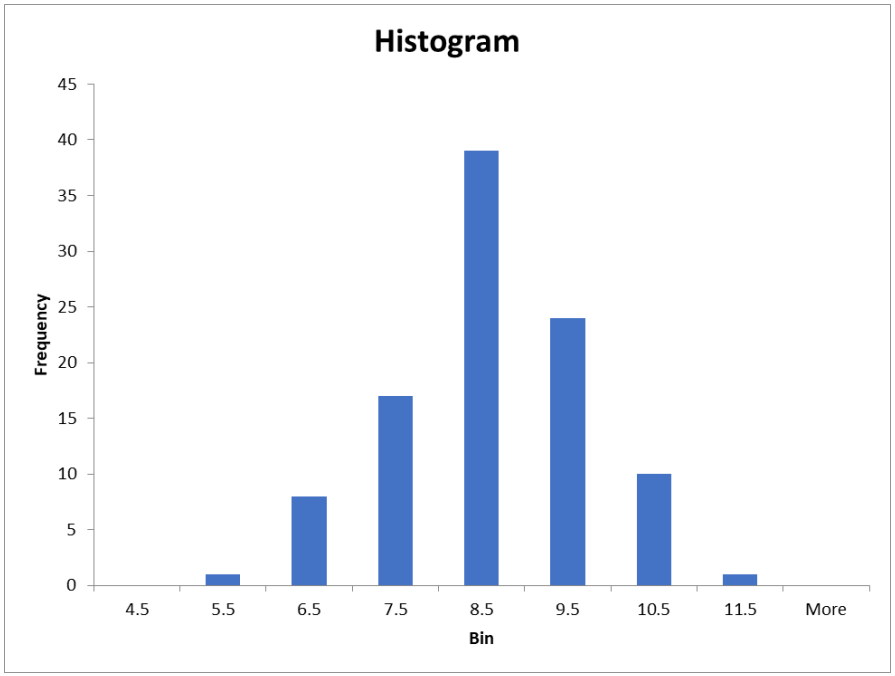
Illustration: A sample of n=100 drawn from a normally distributed population with:

True Mean $\mu = 8$

True SD $\sigma = 1$

Sample Mean $\bar{X}$=8.111

Sample SD $s = 1.038$



2. Calculate Effect Size What is a standard error?

Standard deviation of the sample means

$$SEM = \frac{s}{\sqrt{n}}$$

For the example: $SEM = \frac{1.038}{\sqrt{100}} = 0.104$

Sample Mean $\bar{X}$=8.111 estimates True mean μ = 8.000

with a standard error of 0.104

The standard error tells us about the precision of the estimate.

2. Calculate Effect Size What is a standard error?

Increasing the sample size reduces the standard error of the mean (because the denominator, sqrt(n) is larger).

Increasing the sample size will also reduce the standard error of the effect size.

2. Calculate Effect Size

Standard error for difference in means effect size

$$Difference\ in\ means = \bar{x}_1 - \bar{x}_2$$

$$pooled\ SD, \qquad s = \sqrt{\frac{[(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2]}{(n_1 + n_2 - 2)}}$$

$$Standard\ Error\ for\ Effect\ Size = s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

2. Calculate Effect Size

Standard error for difference in means effect size

We want to weight the studies with a smaller standard error more relative to studies with a larger standard error…

…so we use the **inverse variance weight** or weighting factor (w) which is inverse to the effect size variance (i.e. the square of the standard error)

$$w = \frac{1}{SE^2}$$

2. Calculate Effect Size

Using weighting factors to combine effect sizes

$$\overline{ES} = \frac{\sum(w_i ES_i)}{\sum w_i}$$

The weighted mean effect size can then be calculated by summing the individual weighted effect sizes and dividing by the sum of the weighting factors.

2. Calculate Effect Size

Using weighting factors to combine effects sizes

LF vs LC diets for weight loss

Study 1: (small) n=8

SD=2.9, ES=8.0 kg, SE=1.038 therefore $w_1 = \frac{1}{SE^2} = \frac{1}{1.038^2} = 0.928$

Study 2: (large) n=64

SD=2.9, ES=0.0 kg, SE= 0.363 therefore $w_2 = \frac{1}{SE^2} = \frac{1}{0.363^2} = 7.589$

$$\overline{ES} = \frac{(w_1.ES_1)+(w_2.ES_2)}{(w_1+w_2)} = \frac{(0.928x8)+(7.589x0)}{(0.928+7.589)}$$

$$Summary\ Effect\ Size, \qquad \overline{ES} = 0.87\ kg$$

2. Calculate Effect Size

Example of a different effect size – log Odds Ratio

	Outcome 1	Outcome 2	Odds
Exposure 1	a	b	a/b
Exposure 2	c	d	c/d

$$ES_{LOR} = ln\left(\frac{ad}{bc}\right)$$

$$SE_{LOR} = \sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$$

Weighting factor for Log OR

$$w_{LOR} = \frac{1}{SE^2} = \frac{abcd}{ab(c+d)+cd(a+b)}$$

2. Calculate Effect Size

Studies reporting different types of effect sizes

Research question

Is lead exposure, as measured by plasma lead concentrations, associated with the IQ of children?

1. Study type A

The lead was measured in the plasma of children and their IQ was tested.

Two continuous variables (plasma concentration and IQ)

Effect size – *correlation*

2. Study type B

The lead concentration in the plasma of children was used to classify the children into two groups – lead exposed and unexposed – and the IQ of the children was tested.

One continuous variable (IQ) and one categorical variable (lead exposure)

Effect size – *difference in means*

2. Calculate Effect Size

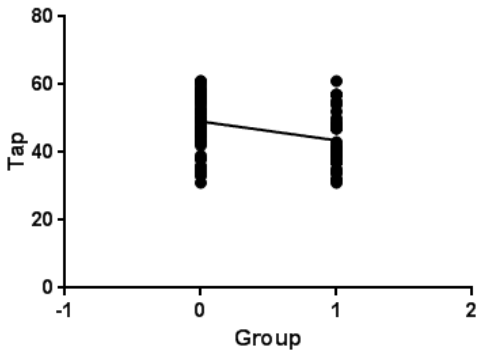
Studies reporting different types of effect sizes

Choices

1. Do analysis on different types of effect sizes separately
2. Use methods to estimate one type of effect size from another

For example:

The Point-Biserial Correlation is a special case of the Pearson's Correlation Coefficient. It is used when one variable is continuous and the other variable is categorical (binomial). For the categorical outcome, one exposure or outcome is assigned 0 and the alternative outcome or exposure is assigned 1.



Example
Group 0 not exposed
Group 1 Lead exposed
Y= IQ score

ES or r = -0.325

2. Calculate Effect Size

Studies reporting different types of effect sizes

2. Use methods to estimate one type of effect size from another

These methods are an estimation and introduce more complexity and uncertainty in determining effect sizes.

For example, a Point-Biserial Correlation is often an underestimate of a correlation coefficient using two continuous variables

Note: Studies that use different measures may differ in other substantive ways. This needs to be considered.

2. Calculate Effect Size Extracting the effect size

In a perfect world…

Standardised effect size for difference in means between two groups, its SE and the weighting factor would be calculated from these values.

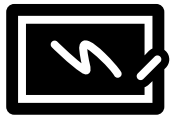
Group	n	Mean	SD
A	34	42.4	6.7
B	28	34.6	2.8

However, it is possible to estimate the standardised effect size from a t statistic or a p value. Eg.

$$Cohen's\ d = t_{(n_A+n_B-2)}\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}$$

Occasionally, non-significant results get reported in insufficient detail for effect size calculation/estimation.

2. Calculate Effect Size Extracting Effect Sizes



Another situation when effect sizes can be difficult to determine is from a multivariate analysis.

The regression model might contain the independent variable of interest, however, its relationship with the dependent variable is adjusted by the presence of other predictor variables in the model.

It is difficult to extract an effect size when only partial regression weights are presented without any kind of univariate analysis.

2. Calculate Effect Size

Estimating the true effect size?

Understanding the (true) mean effect size is assisted by:

1. Weighting effect sizes when calculating a mean effect size.
2. Calculating a confidence interval for the mean effect size.
3. Exploring the heterogeneity of the effect sizes and considering different models for the effect sizes.

2. Calculate Effect Size

Estimating the true effect size?

2. Calculating a confidence interval for the mean effect size.

The 95% Confidence Interval for a mean effect is calculated so there is a 95% probability that this interval contains the population effect size.

It is calculated from the standard error of the mean effect size.

$$SE_{\overline{ES}} = \sqrt{\frac{1}{\sum w_i}}$$

The SE is multiplied by the critical z value representing the desired confidence level (z=1.96 for 95%).

$$95\%\ CI\ for\ \overline{ES} = \overline{ES} \pm 1.96 * SE_{\overline{ES}}$$

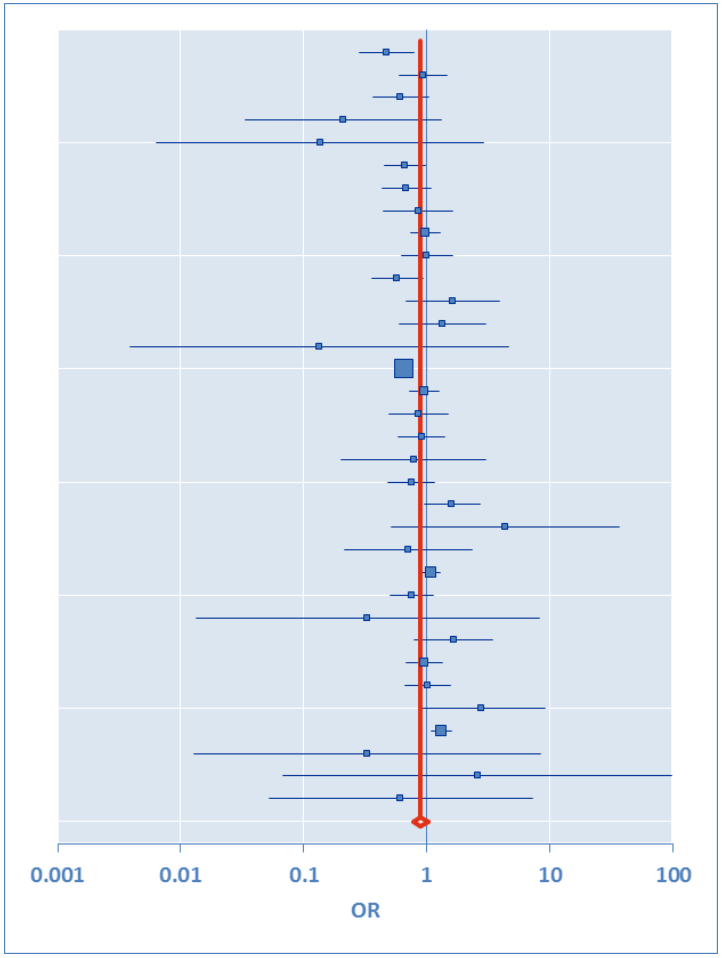
If the 95% CI does not contain 0 then there is a statistically-significant effect at $\alpha = 0.05$.

3. Meta-Analysis summary

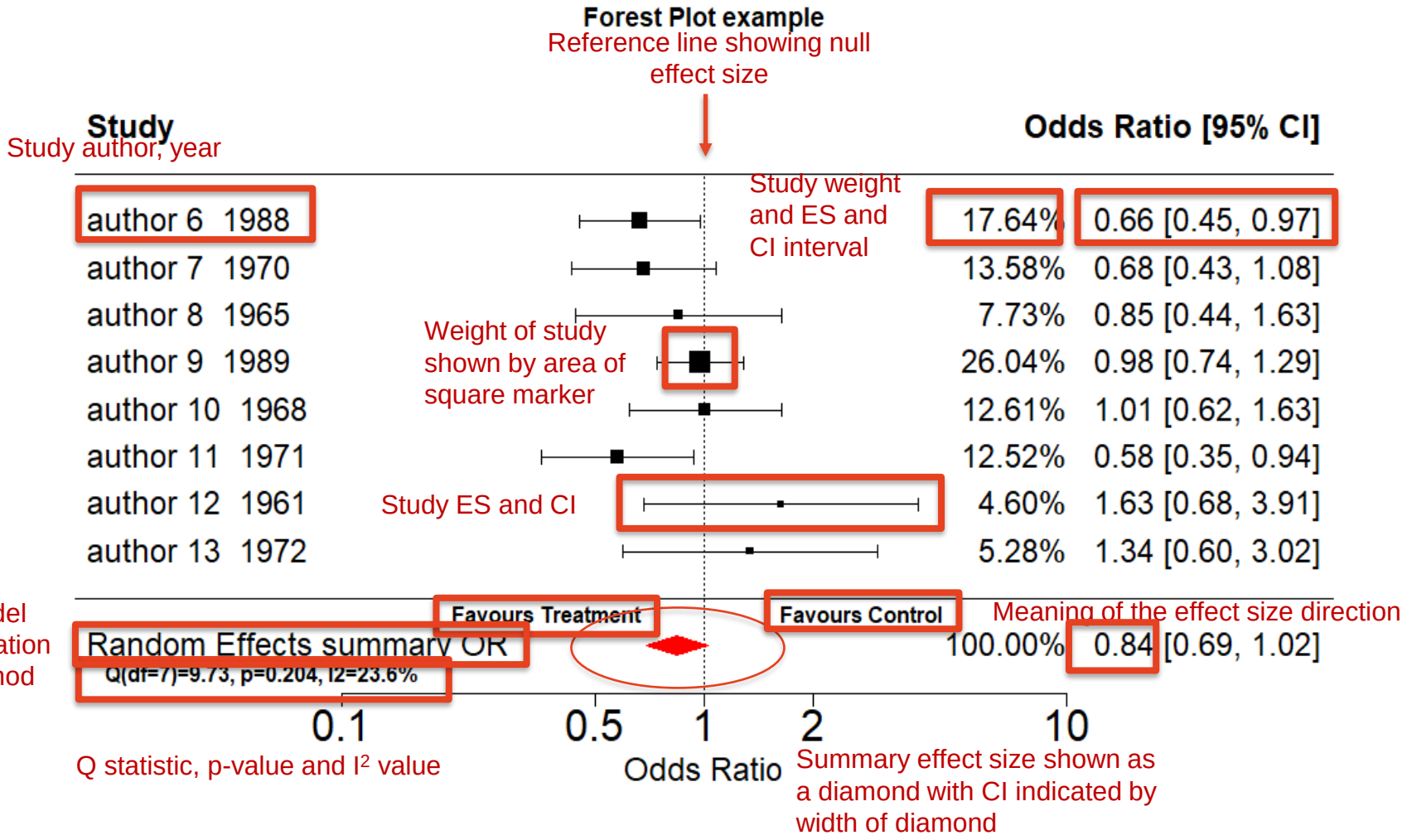
The meta-analysis summary combines all the information we have put together on the study effect sizes, the summary effect size, their associated uncertainties, and information about their spread (or heterogeneity)

The Forest plot is a common tool for displaying this information.

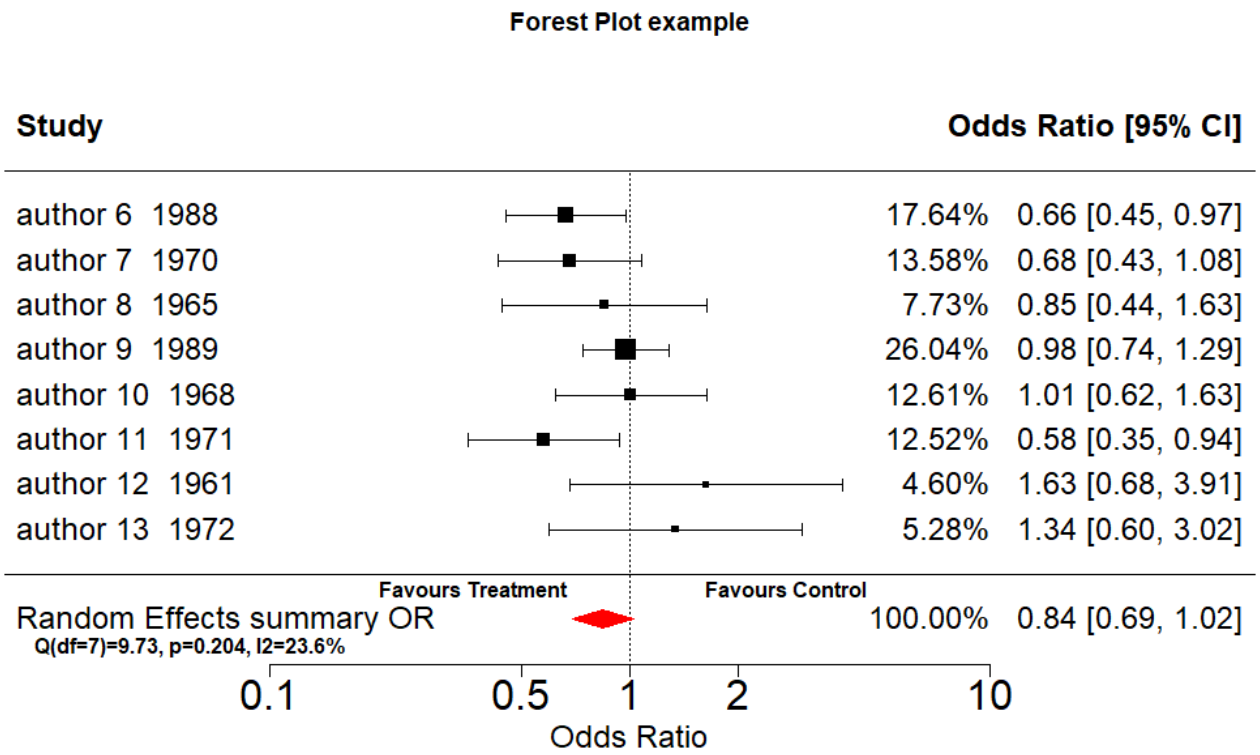
The name refers to the “forest” of lines produced.



3. Meta-Analysis summary Forest Plot



4. Explore Heterogeneity



Do these individual study effect sizes estimate the same underlying true effect size?

4. Explore Heterogeneity Fixed-Effect Models

This model assumes there is one underlying population effect size. The variation in individual effect sizes is due to random sampling.

For example, the population effect of a low fat diet (compared to LC diet) is a weight loss of 0.8 kg.

Due to random sampling error, the individual studies show an effect size ranging from -0.2 to 8.0 kg.

The true effect size is 0.8 kg. If we had infinitely large sample sizes for each study, then we would have ES=0.8 every time, and the standard errors would be zero.

4. Explore Heterogeneity Fixed-Effect Models

Is this a reasonable assumption?

Why?

4. Explore Heterogeneity

Assumption of Homogeneity – Q statistic

Is it reasonable to assume the distribution of observed effect sizes due to random sampling is from a population with a single effect size?

The homogeneity test is based on the (Cochran's) Q statistic

$$Q = \sum_{i=1}^{k} w_i (ES_i - \overline{ES})^2$$

The Q statistic has a chi-square distribution with *k-1* degrees of freedom where *k* is the number of effect sizes.

If the observed Q statistic is over the critical value (producing a P value less than α) then the null hypothesis (homogeneity of effect sizes) is rejected.

4. Explore Heterogeneity

Quantifying heterogeneity with the I^2 statistic

Effect sizes are always going to be heterogeneous. Why not quantify the inconsistency rather than test with some arbitrary cut-off?

A common approach is to apportion the amount of variation due to heterogeneity in study effect sizes and the amount of variation due to sampling error.

$$I^2 = 100\% \times \left(\frac{Q - df}{Q} \right)$$

I^2 conveys the amount of variation in the effects sizes that would remain if we could remove sampling error.

4. Explore Heterogeneity

Quantifying heterogeneity with the I^2 statistic

What is a large I^2 value?

Rules of thumb:

from Borenstein:

- 25% is low
- 50% is moderate
- 75% is high

From Cochrane Training:

- 0% to 40%: might not be important
- 30% to 60%: may represent moderate heterogeneity
- 50% to 90%: may represent substantial heterogeneity
- 75% to 100%: considerable heterogeneity

4. Explore Heterogeneity Random Effects models

This model assumes that there are differences in the underlying effect sizes whose sources cannot be identified.

Therefore the variation in effects sizes is due to sampling error (as with fixed effects models) but with an additional component that estimates other sources of variability.

$$v_{RE} = v_\theta + v_i$$

Where

v_i is the estimate of the variance associated with sampling error.

v_θ is the estimate of the between-studies variance, also known as tau-squared (τ^2).

4. Explore Heterogeneity Random Effects models

v_θ is the estimate of the between-studies variance, also known as tau-squared (τ^2).

The calculation of τ^2 can get technical. It is common to use the method of moments known as the "DerSimonian and Laird" method.

In the workshop example we will use the REML estimate.

4. Explore Heterogeneity

The consequences of using a Random Effects Model

The inverse variance weights calculated previously (and thus the mean effect sizes) assumed the fixed effects model.

$$inverse\ variance\ weight = \frac{1}{SE^2}$$

However, with the random effects model that becomes

$$inverse\ variance\ weight = \frac{1}{v_{RE}}$$

Therefore, when switching to a random effects model, all mean effect sizes and confidence intervals are recalculated.

4. Explore Heterogeneity

A commonly used approach

Approach the analysis using the fixed effect model as the default model and explore the heterogeneity of the effects sizes using a statistical test (Q statistic).

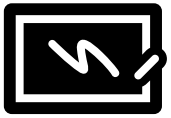
1. The test for heterogeneity is not significant therefore a fixed effect model is consistent with the data and this model is used; OR

2. The test for heterogeneity is significant so the analysis is repeated with the random effects model.

The problem with this approach…

The non-significant p value does not prove the null hypothesis (that there is one true effect size).

It is recommended that the decision is based on rigorous assumptions regarding the sample of studies.



How should we select a statistical model?

The choice of a statistical model should depend on the sampling frame that was used to select studies for the analysis. If we are working with one population, then we should use the fixed-effect model. If we are working with a universe of populations, we should use the random-effects model. Consider the following two cases.

Case 1. A pharmaceutical company plans to enroll 1,000 patients. Because the staff can only work with 100 patients at time, it randomly divides the patients into 10 groups of 100 patients, and runs the identical study with each. We know all studies are based on the same population, since that's how we selected them. The fixed-effect model matches the way the studies were sampled.

Case 2. We locate a series of published studies that were performed by different people in different locations at different times. While the studies all looked at similar interventions, it stands to reason that the true impact of this intervention will differ from study to study. If the studies were conducted in different hospitals it's likely that the patient population (age, co-morbid diseases) varied from one hospital to the next, and that the intervention was therefore more effective in some hospitals than in others. It's also possible that the intervention itself (the precise dosage, length of follow-up) differed from one hospital to the next and that this could have an impact on the effect size.

From the metaanalysis.com website

4. Explore Heterogeneity Prospective Meta-analysis

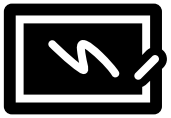
Cochrane Methods

"A prospective meta-analysis (PMA) is a meta-analysis in which studies (usually randomised controlled trials) are identified, evaluated and determined to be eligible before the results of any of the studies become known. PMA can help to overcome some of the recognised problems of retrospective meta-analyses."

- This approach may lead to low heterogeneity and a justifiable assumption of Fixed Effects

4. Explore Heterogeneity

Fixed-Effect and Random-Effects summary



- A fixed-effect meta-analysis estimates a single effect that is assumed to be common to every study, while a random-effects meta-analysis estimates the mean of a distribution of effects.
- Study weights are more balanced under the random-effects model than under the fixed-effect model. Large studies are assigned less relative weight and small studies are assigned more relative weight as compared with the fixed-effect model.
- The standard error of the summary effect and (it follows) the confidence intervals for the summary effect are wider under the random-effects model than under the fixed-effect model.
- The selection of a model must be based solely on the question of which model fits the distribution of effect sizes, and takes account of the relevant source(s) of error. When studies are gathered from the published literature, the random effects model is generally a more plausible match.
- The strategy of starting with a fixed-effect model and then moving to a random-effects model if the test for heterogeneity is significant is a mistake, and should be strongly discouraged.

Borenstein et. al., 2009

4. Explore Heterogeneity Selecting a model

A random effects model would fit most types of meta-analyses, based on the sources of studies.

At first glance, the random effects model would appear to be the more conservative approach.

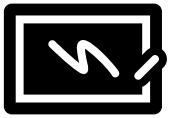
However…

Random effects incorporate the between study variance into the denominator of each weight. This has the effect of weighting the studies more evenly and can lead to the effect size being shifted towards the effect size of smaller studies.

Perhaps not a problem unless there is a different effect size observed among the smaller studies compared to the larger studies.

How can this be assessed?

5. Publication Bias



Selection bias

Only choose studies which they are familiar with or support their view-point

Occurs when the two groups being compared differ systematically, eg. differences in the characteristics between those who are selected for a study.

Performance bias

differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest.

Detection bias

differences between groups in how outcomes are determined.

Attrition bias

differences between groups in withdrawals from a study.

Reporting bias

differences between reported and unreported findings.

Publication bias

Only select studies which have been published. In general, studies with statistically significant or positive results are more likely to be published than those with non-significant or negative results.

Language bias

Only select studies written in English

Indexing bias

Only search a limited number of databases or use a limited choice of search terms.

5. Publication Bias

"The small proportion of results chosen for publication are unrepresentative of scientists' repeated samplings of the real world."

Young, Ioannidis & Al-Ubaydli, PL0S Medicine, doi:10.1371

"The winner's curse refers to the phenomenon that studies that find evidence of an effect often provide inflated estimates of the size of that effect. Such inflation is expected when an effect has to pass a certain threshold — such as reaching statistical significance — in order for it to have been 'discovered'. Effect inflation is worst for small, low-powered studies, which can only detect effects that happen to be large."

Button et al, Nature Reviews Neuroscience, 2013,14:365

5. Publication Bias what can you do?

- As individuals, we can't change the forces that cause publication bias.
- Look at studies that are available from other sources and may be more representative.
- Research Grey Literature sources
- Avoid indexing bias! (a more thorough literature search will generate a more valid meta-analysis)

5. Publication Bias Funnel Plot

A funnel plot is a scatter plot of each study's Effect Size vs a measure of the study's size or precision e.g. SE of the effect size.

The next few figures are taken from Sterne et al., 2011;342:d4002

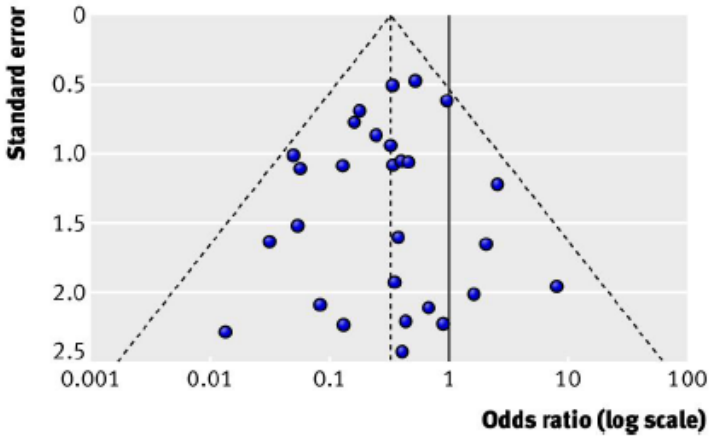


Fig 1 Example of symmetrical funnel plot. The outer dashed lines indicate the triangular region within which 95% of studies are expected to lie in the absence of both biases and heterogeneity (fixed effect summary log odds ratio±1.96×standard error of summary log odds ratio). The solid vertical line corresponds to no intervention effect

5. Publication Bias Funnel Plot Asymmetry

It has been observed that small studies report larger effect sizes compared with large studies. This may be due to publication bias but could also be due to other factors e.g. different subgroups in the smaller studies.

A different effect size in the small studies can be visualised as asymmetry in the funnel plot.

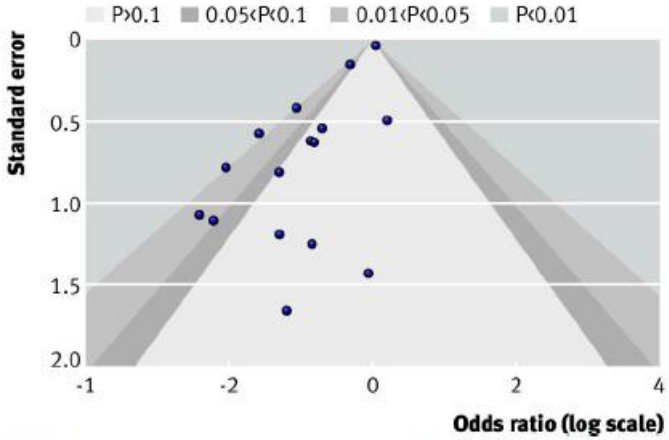


Fig 4 Contour enhanced funnel plot for trials of the effect of intravenous magnesium on mortality after myocardial infarction

5. Publication Bias Test for funnel plot asymmetry

It is possible to test for asymmetry but check the recommendations as it may not be advisable.

For example,

- The minimum number of studies should be substantially greater than 10.
- The asymmetry test should not be used if the studies are of similar size/precision.
- The plots need to be considered in the context of the other characteristics of the studies.

If there is significant asymmetry, the effect estimate should be compared under the fixed and random-effects model as the random effects model might be skewing the mean effect size towards the smaller studies.

5. Publication Bias Funnel Plot Asymmetry

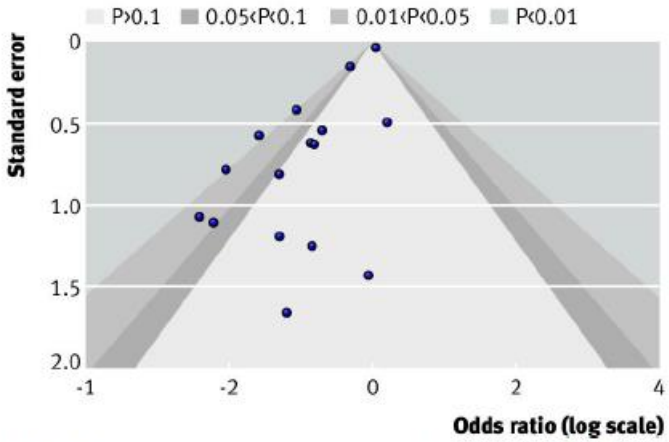


Fig 4 Contour enhanced funnel plot for trials of the effect of intravenous magnesium on mortality after myocardial infarction

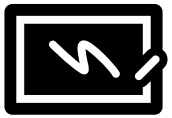
One very large study showing no beneficial effect of intervention and other small studies showing mostly beneficial effect.

Fixed Effects model – No net effect

Random Effects model – An overall beneficial effect

This is because the missing study data (RHS of plot) results in a smaller between study variance than should be the case.

5. Publication Bias Other diagnostics



Duval and Tweedie – Trim and Fill (for a funnel plot)

Cumulative Forest plot

Egger's test for asymmetry of funnel plot

Other tests like Egger's (Begg's, Deeks', etc)

6. Sub-group analysis moderating factor

1. There is one underlying population effect size. The variation in individual effect sizes is due to random sampling.

For example, the population effect of a low fat diet (compared to LC diet) is a weight loss of 0.8 kg.

Due to random sampling error, the individual studies show an effect size ranging from -0.2 to 8.0 kg.

OR

2. There is variation in the population effect sizes but this is systematic and due to a moderating factor.

For example, the population effect of the LF diet on weight loss is 1.5 kg in older adults whilst the effect is 0.5 kg in younger adults.

6. Sub-group analysis moderating factor

If we suspect a moderating factor is present (and we have the data for it)…

…then the homogeneity analysis can be partitioned according to the grouping factor.

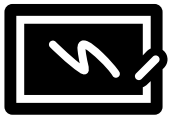
For example,

The categorical or grouping variable is Age Group (older adults or younger adults).

The Q statistic is partitioned into the portion that is explained by Age Group (Q_B) and the residual portion that explains the within group variation (Q_W).

$$Q = Q_B + Q_W$$

6. Sub-group analysis Meta-regression



Meta-regression is a more flexible approach to understanding the factors that might contribute to an effect size because:

- The moderator can be a numeric variable instead of being restricted to a categorical variable.
- Multiple variables (including categorical variables when coded appropriately) can be considered in the same analysis.

Example: Weight loss for LF diet (vs LC diet)

Rather than compare sub-groups for younger vs older adults, if we had information on subject age on a continuous scale then we could incorporate this covariate in a meta-regression model.

Like all multiple regression models, a large ratio of subjects (i.e. studies) to covariates is required (suggestion of 10:1).

The regression model can be formulated using the assumptions of the fixed or random effects model.

Meta-Analysis workflow

So just to re-cap, here are the basic workflow steps we went through:

1. Import the data
2. Calculate Effect Size
3. Meta-Analysis summary (Fixed-Effect and Random-Effects)
4. Explore Heterogeneity (Q statistic, I^2, tau)
5. Publication bias (Funnel plot)
6. Sub-group analysis and/or meta-regression (if warranted)

Meta Analysis software

Name, URL	Comment	Interface
SPSS https://mason.gmu.edu/~dwilsonb/ma.html	3rd party macro	SPSS syntax only
STATA https://www.stata.com/new-in-stata/meta-analysis/	From version 16	Command meta
CMA www.meta-analysis.com	Developed by Borenstein (world expert). Costs $$$. Popular	GUI
MIX https://www.meta-analysis-made-easy.com/	Add-In for Excel (Windows), user friendly, Modest $	Excel menus
Revman https://training.cochrane.org/online-learning/core-software-cochrane-reviews/revman	Open-source (free) from Cochrane Training Web and Desktop versions	GUI

There are of course other options not listed here…

Meta Analysis in R

Software or package	Origin	Comments
meta	Schwarzer, 2007	General package
rmeta	Lumley, 2009	General package
dmetar	Harrer, et al	Can be used with meta and metafor
metap	Dewey, 2020	Enables meta-analysis using significance values
metafor	Viechtbauer, 2010	has more features than meta or rmeta
jamovi https://www.jamovi.org/	uses metafor R package	User friendly interface, can save R script
mada	Doebler, 2020	Meta-Analysis of Diagnostic Accuracy (for summary ROC)

There will also no doubt be other packages not listed here…

The metafor package

www.metafor-project.org/doku.php

The metafor Package
A Meta-Analysis Package for R

Search

Media Mana

metafor

avigation

- Homepage
- Package News
- Package Features
- Package Update Log
- To-Do List / Planned Features
- Download and Installation
- Documentation and Help
- Analysis Examples
- Plots and Figures
- Tips and Notes
- Contributors
- FAQs
- Links

xternal Links

- Wolfgang Viechtbauer
- The R Project
- CRAN

The metafor Package: A Meta-Analysis Package for R

The metafor package is a free and open-source add-on for conducting meta-analyses with the statistical software environment R. The package consists of a collection of functions that allow the user to calculate various effect size or outcome measures, fit fixed-, random-, and mixed-effects models to such data, carry out moderator and meta-regression analyses, and create various types of meta-analytical plots.

On this website, you can find:

- some news concerning the package and/or its development,
- a more detailed description of the package features,
- a log of the package updates that have been made over the years,
- a to-do list and a description of planned features to be implemented in the future,
- information on how to download and install the package,
- information on how to obtain documentation and help with using the package,
- some analysis examples that illustrate various models, methods, and techniques,
- a little showcase of plots and figures that can be created with the package,
- some tips and notes that may be useful when working with the package,
- a list of people that have in some shape or form contributed to the development of the package,
- a frequently asked questions section, and
- some links to other websites related to software for meta-analysis.

The metafor package was written by Wolfgang Viechtbauer. It is licensed under the GNU General Public License Version 2. For citation info, type `citation(package='metafor')` in R. To report any issues or bugs, please go here.

metafor.txt · Last modified: 2019/02/06 22:16 (external edit)

The metafor package

last updated: May 09 2018
(not all functions documented)

given the required data (e.g., means, SDs, and group sizes; counts for 2x2 tables; correlations and sample sizes), calculate the desired effect size or outcome measure for the meta-analysis (e.g., raw or standardized mean differences, log odds ratios, log risk ratios, risk differences, r-to-z transformed correlations, ...)

read.table()
read.csv()
read.delim()

functions in the 'util' package to:
- read in data from ASCII file
- see also 'foreign' package for reading in other data formats

- yi = observed outcomes or effect size estimates
- vi = corresponding sampling variances

rma.uni()
rma.mh()
rma.peto()
rma.glmm()
rma.mv()

- rma.uni() = fixed- and random/mixed-effects models ("inverse-variance" method; normal-normal models)
- rma.mh() = Mantel-Haenszel method (fixed-effects model)
- rma.peto() = Peto's method (fixed-effects model)
- rma.glmm() = fixed- and random/mixed-effects models (binomial-normal and Poisson-normal models)
- rma.mv() = fixed- and random/mixed-effects multivariate/multilevel models (normal-normal models)

note: rma.uni() takes either 'yi' and 'vi' as input or one can supply the required data to calculate the desired effect size or outcome measure for the meta-analysis directly; rma.mh(), rma.peto(), and rma.glmm() require that the raw counts are supplied; rma.mv() takes 'yi' and 'V' as input (V is the variance-covariance matrix of the sampling errors)

print()
summary()

print functions

print()
summary()

note: class of fitted model object is the same as the function name; so print() for an object of class 'rma.uni' actually calls print.rma.uni() and so on

fitted and predicted values

fitted()
predict()
blup()
ranef()
cumul()

note: blup() only for 'rma.uni' objects; ranef() only for 'rma.uni' and 'rma.mv' objects; cumul() not for 'rma.mv' or 'rma.glmm' objects

residuals and influential case diagnostics

residuals()
rstandard()
rstudent()
hatvalues()
weights()
influence()
leave1out()

note: all functions implemented for 'rma.uni' objects; coverage of functions for other objects is more limited (see docs)

funnel plot asymmetry (publication bias)

ranktest()
regtest()
trimfill()
hc()

note: regtest() not for 'rma.glmm' or 'rma.mv' objects; trimfill() and hc() only for 'rma.uni' objects

confidence intervals and inference

confint()
anova()
permutest()
robust()

note: confint() not for 'rma.glmm' objects; anova() and robust() only for 'rma.uni' and 'rma.mv' objects; permutest() only for 'rma.uni' objects

plotting functions

forest()
funnel()
labbe()
radial()
qqnorm()
baujat()
gosh()
plot()

note: forest() and funnel() also take 'yi' and 'vi' as input; qqnorm(), baujat(), gosh() and plot() not for 'rma.glmm' or 'rma.mv' objects

various extractor functions

logLik()
deviance()
fitstats()
AIC(), BIC()
coef()
vcov()

note: coef() also for 'permutest.rma.uni' and 'summary.rma' objects

Example: Meta analysis with metafor in R

Sutton-Smith Cholesterol lowering intervention on mortality

This data is taken from "Methods of Meta-analysis in Medical Research" by A. J. Sutton (originally from *British Medical Journal* 1993;306:1367 published as "Cholesterol lowering and mortality: the importance of considering initial level of risk")

The data comprises 34 studies assessing the effectiveness of cholesterol treatments on mortality in randomised controlled trials.

Example: Meta analysis with metafor in R

Sutton-Smith Cholesterol lowering intervention on mortality

The following files are provided for you to follow this example on your own device:

The data: Meta_Sutton_Smith.csv

The R script: Meta-Analysis workshop example.R

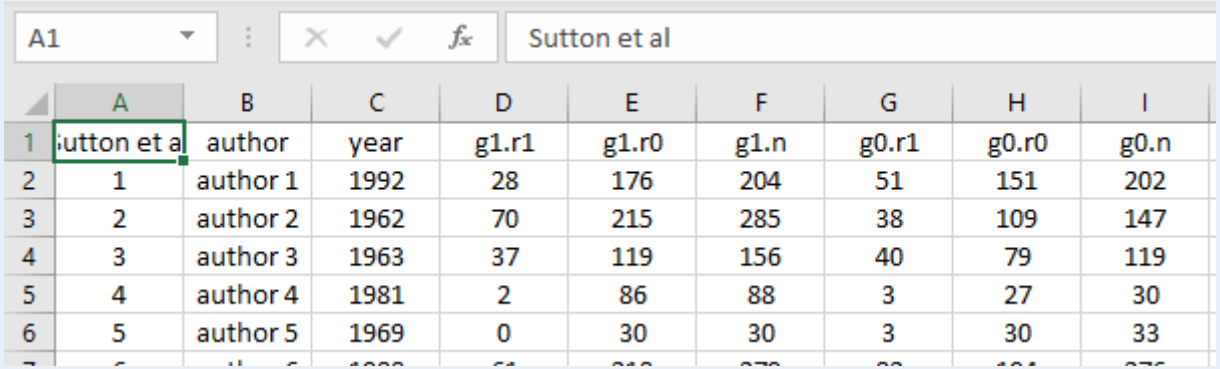
If you are not an R user, just follow the output. A similar workflow can be followed using other software tools.

Example: Sutton-Smith (binary outcome)

Step 1: Import the data

We have the data in a csv file ready to read in using:

```
metadata <- read.csv("Meta_Sutton_Smith.csv", header = T)
```



A1 | Sutton et al

	A	B	C	D	E	F	G	H	I
1	Sutton et al	author	year	g1.r1	g1.r0	g1.n	g0.r1	g0.r0	g0.n
2	1	author 1	1992	28	176	204	51	151	202
3	2	author 2	1962	70	215	285	38	109	147
4	3	author 3	1963	37	119	156	40	79	119
5	4	author 4	1981	2	86	88	3	27	30
6	5	author 5	1969	0	30	30	3	30	33

Example: Sutton-Smith (binary outcome)

Step 1: Import the data

These studies compare the mortality rates for Group 1 (intervention) and Group 0 (control).

The data is presented in counts where r1=event, r0=no event, n=total

				Group 1			Group 0		
	Sutton.et.al	author	year	g1.r1	g1.r0	g1.n	g0.r1	g0.r0	g0.n
1	1	author 1	1992	28	176	204	51	151	202
2	2	author 2	1962	70	215	285	38	109	147
3	3	author 3	1963	37	119	156	40	79	119
4	4	author 4	1981	2	86	88	3	27	30
5	5	author 5	1969	0	30	30	3	30	33
6	6	author 6	1988	61	218	279	82	194	276
7	7	author 7	1970	41	165	206	55	151	206
8	8	author 8	1965	20	103	123	24	105	129
9	9	author 9	1989	111	907	1018	113	902	1015
10	10	author 10	1968	81	346	427	27	116	143

Example: Sutton-Smith (binary outcome)

Step 2: Calculate effect size (Odds Ratios)

We use the "escalc" function from metafor to create "yi" which is the OR effect size and "vi" which is the variance of "yi".

```
meta.es <- escalc(measure = "OR", g1.r1, g1.r0, g0.r1, g0.r0,
data = metadata, slab = paste(author, year, sep = ", "))
```

	Sutton.et.al	author	year	g1.r1	g1.r0	g1.n	g0.r1	g0.r0	g0.n	yi	vi
1	1	author 1	1992	28	176	204	51	151	202	-0.752825281	0.067626464
2	2	author 2	1962	70	215	285	38	109	147	-0.068381064	0.054426978
3	3	author 3	1963	37	119	156	40	79	119	-0.487637182	0.073088616
4	4	author 4	1981	2	86	88	3	27	30	-1.563975538	0.881998277

Example: Sutton-Smith (binary outcome)

Step 3: Calculate meta-analysis summary

We use the “rma” function from metafor to create the meta analysis summary statistics

Random effects is the default

```
meta.re <- rma(yi, vi, data = meta.es)
```

Fixed effect requires an optional argument

```
meta.fe <- rma(yi, vi, data = meta.es, method = "FE")
```

Example: Sutton-Smith (binary outcome)

Step 3: Calculate meta-analysis summary

Random-Effects model

```
Random-Effects Model (k = 34; tau^2 estimator: REML)

tau^2 (estimated amount of total heterogeneity): 0.0478 (SE = 0.0269)
tau (square root of estimated tau^2 value):      0.2186
I^2 (total heterogeneity / total variability):   53.87%
H^2 (total variability / sampling variability):  2.17

Test for Heterogeneity:
Q(df = 33) = 89.1065, p-val < .0001

Model Results:

estimate      se     zval    pval    ci.lb   ci.ub
 -0.1178  0.0628  -1.8750  0.0608  -0.2410  0.0053   .

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note default REML estimation

Information on the summary effect size as log(OR)

Example: Sutton-Smith (binary outcome)

Step 3: Calculate meta-analysis summary

Fixed-Effect model

```
Fixed-Effects Model (k = 34)

I^2 (total heterogeneity / total variability):   62.97%
H^2 (total variability / sampling variability):  2.70

Test for Heterogeneity:
Q(df = 33) = 89.1065, p-val < .0001

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.1691  0.0324  -5.2258  <.0001  -0.2325  -0.1057  ***

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Information on the summary effect size as log(OR)

Example: Sutton-Smith (binary outcome)

Step 3: Calculate meta-analysis summary

Model comparison (convert log(OR) to OR first)

```
meta.re.or <- exp(c(meta.re$b, meta.re$ci.lb, meta.re$ci.ub))
meta.fe.or <- exp(c(meta.fe$b, meta.fe$ci.lb, meta.fe$ci.ub))
```

Model	Effect size	lower CI	Upper CI	p value
Random-Effects	0.889	0.786	1.005	0.0608
Fixed-Effect	0.844	0.793	0.900	<.0001

Random-Effects model has

- Reduced OR Effect size from 0.84 to 0.89 (closer to 1)
- Increased width of CI from 0.79~0.90 to 0.79~1.01
- CI for effect size now traverses 1.00, p value >0.05

Example: Sutton-Smith (binary outcome)

Step 3: Calculate meta-analysis summary

Forest Plot – R code

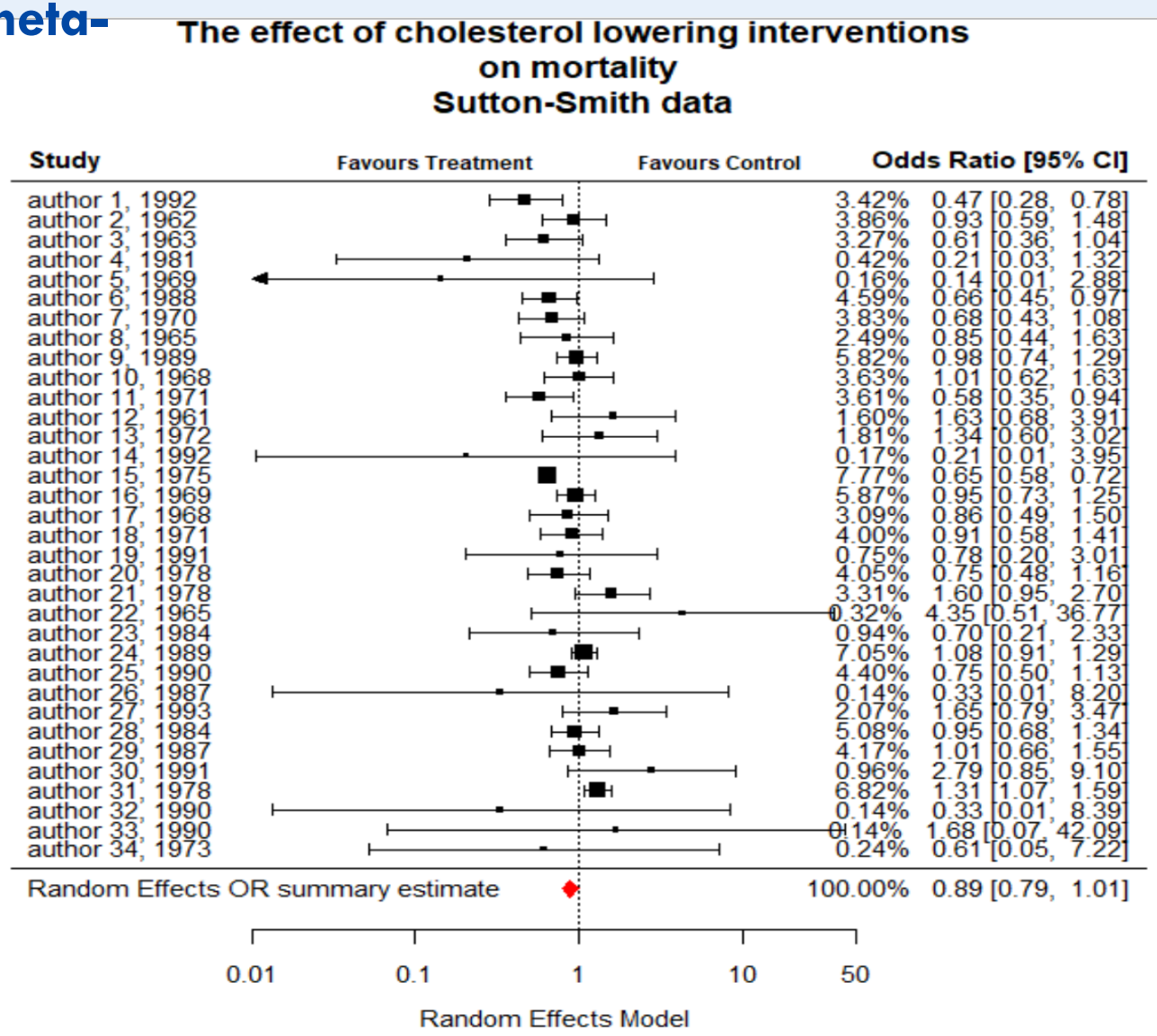
Use the forest() function to produce the plot. You may need to adjust quite a few plotting parameters to get it just the way you want. This is the code I used for the fixed-effect forest plot (code for random not shown).

```
forest(
  meta.fe,
  header = TRUE,
  atransf = exp, #display OR on x-axis log scale
  at = log(c(0.01, 0.1, 1, 10, 50)),
  xlim = c(-8, 8),
  xlab = "Fixed-Effect Model",
  mlab = "Fixed-Effect OR summary estimate",
  showweights = TRUE,
  pch = 15,
  border = "red",  #colour of summary diamond border
  col = "red",     #colour of summary diamond fill
  main = "The effect of cholesterol lowering interventions \n on mortality \n Sutton-Smith data"
)
op <- par(cex = 0.75, font = 2)
text(c(-2, 2), 36, c("Favours Treatment", "Favours Control"))
par(op)
```

Example: Sutton-Smith (binary outcome)

Step 3: Calculate meta-analysis summary

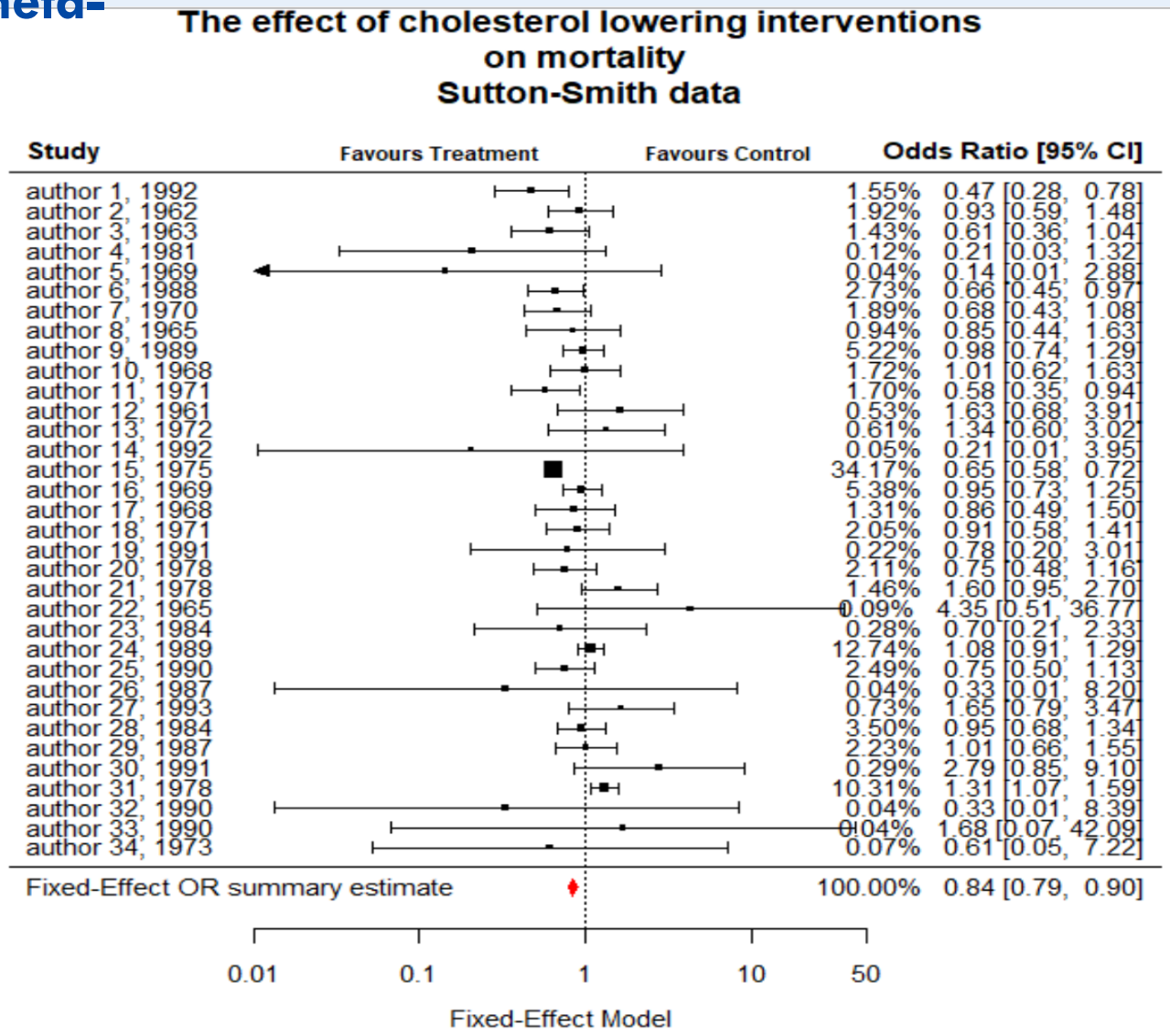
Forest Plot – Random-Effects



Example: Sutton-Smith (binary outcome)

Step 3: Calculate meta-analysis summary

Forest Plot – Fixed-Effect



Example: Sutton-Smith (binary outcome)

Step 3: Calculate meta-analysis summary

Forest Plot

Comments

- Note estimates and CI's for each study. Each study's weight designated by size of square.
- Graphical representation of (moderate) heterogeneity.
- Weightings are different between the two models

Example: Sutton-Smith (binary outcome)

Step 4: Explore heterogeneity

- Q statistic is significant $p<0.0001$ (heterogeneous studies)
- I squared (proportion of inconsistency) 54%~63% is moderate
- Tau squared estimate is > zero
- We expect studies to have underlying differences in effect size
- We should use random effects

Example: Sutton-Smith (binary outcome)

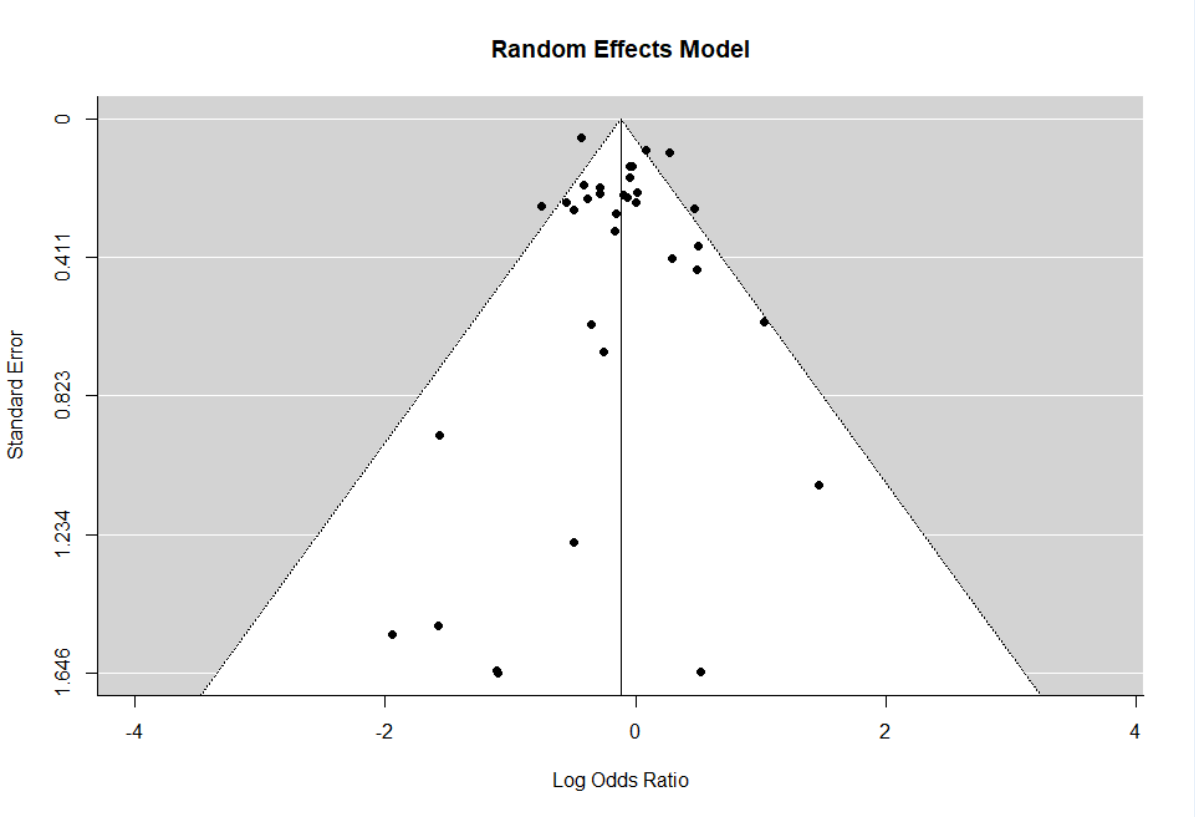
Step 5: Publication Bias

- Funnel Plot

funnel(meta.re, main = "Random Effects Model")

Check for asymmetry

- No of studies > 10
- Studies are of various sizes and precision
- Plot appears to be symmetrical (can use random effects model without problem)
- One more check



Example: Sutton-Smith (binary outcome)

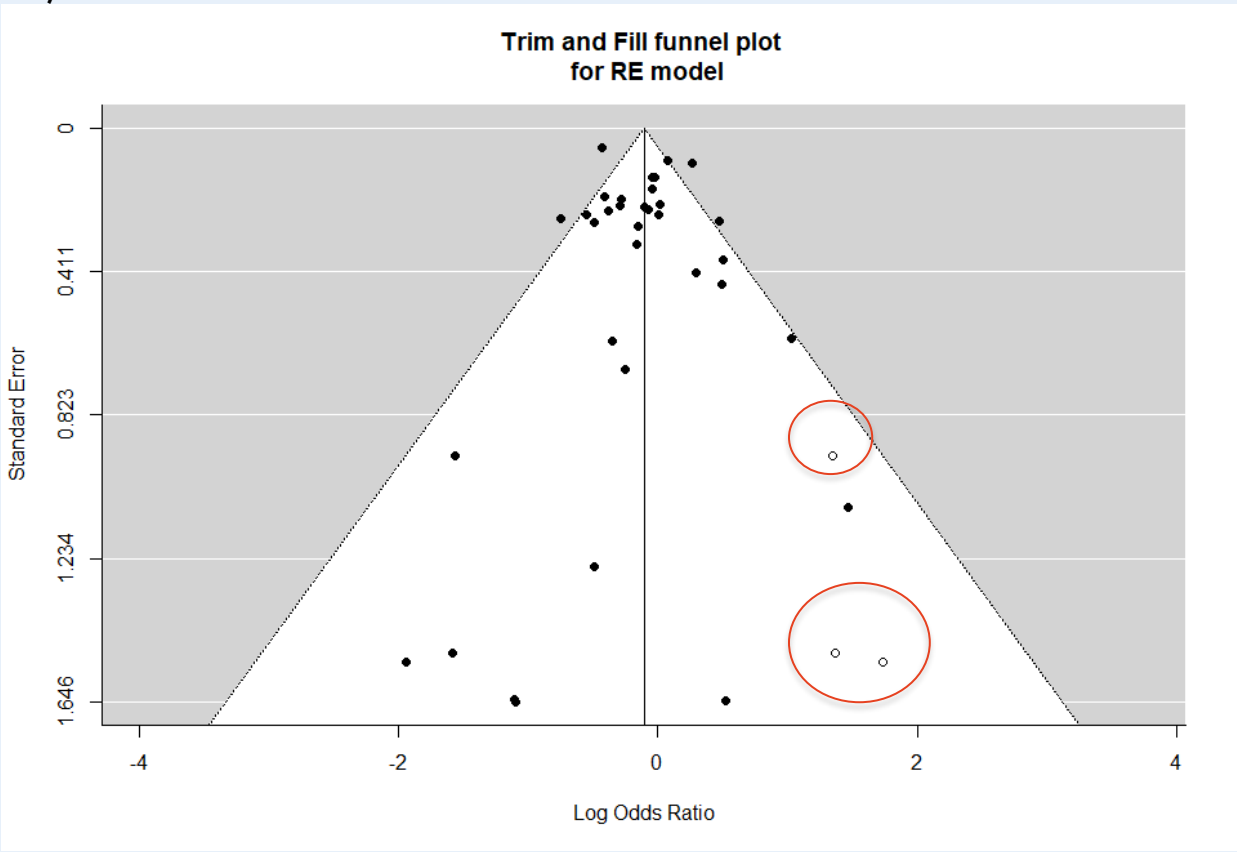
Step 5: Publication Bias

Funnel Plot

Use trimfill() function to add imputed studies to the funnel plot

```
meta.tf <- trimfill(meta.re)
funnel(meta.tf, main = "Trim and Fill funnel plot \n for RE model")
```

Note: 3 imputed studies (open circles)



Example: Sutton-Smith (binary outcome)

Step 6: Sub-group analysis

The original study looked at a sub-group analysis using the factor:

Coronary heart disease risk group – high, medium, low

Sub-group analysis found significant differences between these groups, with only the high risk CHD (treatment) group showing a benefit in terms of total mortality.

A meta-regression was also carried out showing a similar trend.

(This analysis is not included in this workshop)

Published Meta-analysis examples

1. When publication bias is present – from Experimental Psychology (inspect the funnel plot)

Shanks, David R et al. "Romance, Risk, and Replication: Can Consumer Choices and Risk-Taking Be Primed by Mating Motives?" Journal of experimental psychology. General 144.6 (2015): e142–e158. Web.

https://sydney.primo.exlibrisgroup.com/permalink/61USYD_INST/1367smt/cdi_gale_infotracacademiconefile_A444328766

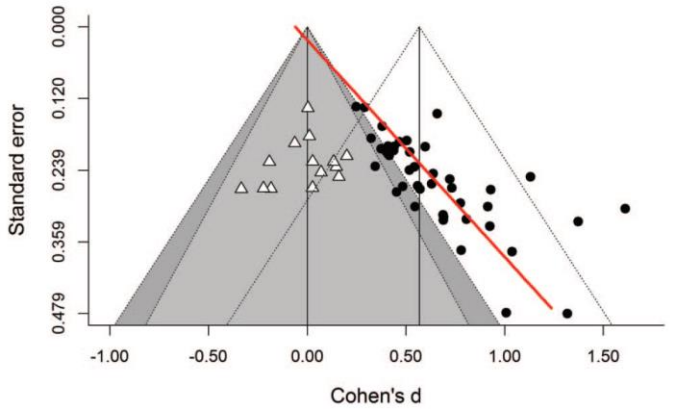


Figure 2. Funnel plot from the meta-analysis. Black circles represent the

2. Good example of tabulation of study data (Tables 1,2,3) but bad example of plotting meta-regression (figures 2 and 4)

Rong, Chen. "Egg Consumption and Risk of Coronary Heart Disease and Stroke: Dose-Response Meta-Analysis of Prospective Cohort Studies." BMJ : British Medical Journal 346.7890 (2013): 12–12. Web.

https://sydney.primo.exlibrisgroup.com/permalink/61USYD_INST/1367smt/cdi_proquest_miscellaneous_1273301943

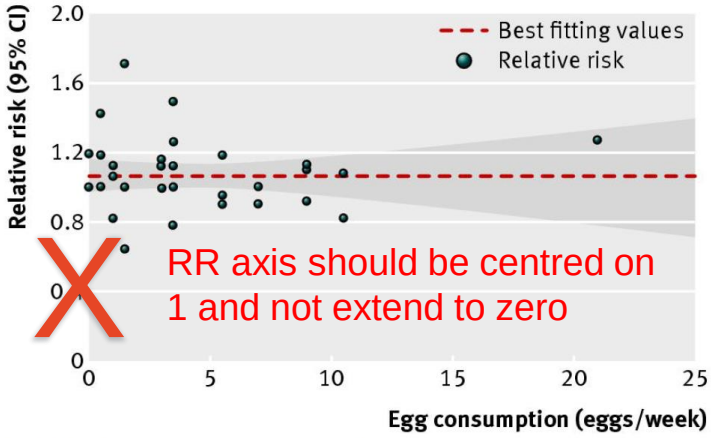


Fig 2 Dose-response analyses of egg consumption and

Published Meta-analysis examples

3. Scientific fraud and retracted papers examined by sub-group analysis

Zarychanski, Ryan et al. "Association of Hydroxyethyl Starch Administration With Mortality and Acute Kidney Injury in Critically Ill Patients Requiring Volume Resuscitation: A Systematic Review and Meta-Analysis." JAMA : the journal of the American Medical Association 309.7 (2013): 678–688. Web.

https://sydney.primo.exlibrisgroup.com/permalink/61USYD_INST/1367smt/cdi_proquest_miscellaneous_1291608126

4. Good example of Cumulative meta-analysis over 17yr

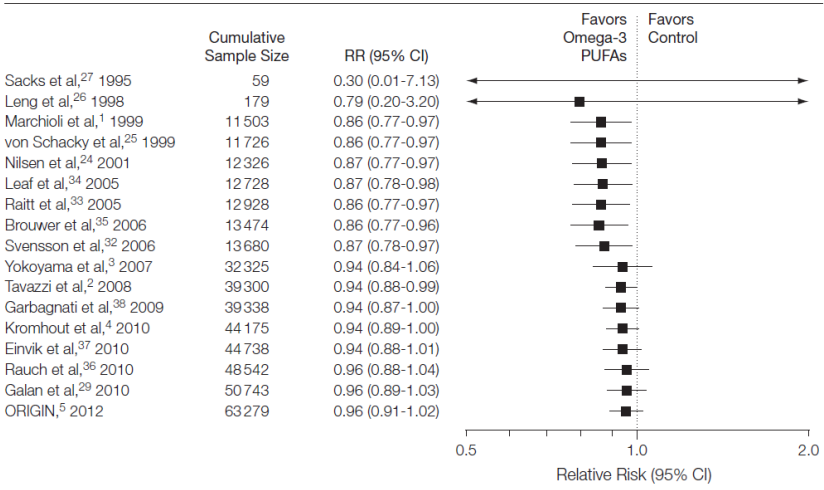
Rizos, Ntzani. "Association Between Omega-3 Fatty Acid Supplementation and Risk of Major Cardiovascular Disease Events: A Systematic Review and Meta-Analysis." JAMA : the journal of the American Medical Association 308.10 (2012): 1024–1033. Web.

https://sydney.primo.exlibrisgroup.com/permalink/61USYD_INST/1367smt/cdi_proquest_miscellaneous_1039344012

Retraction Watch

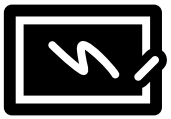
Tracking retractions as a window into the scientific process

Figure 5. Cumulative Meta-analysis of the Omega-3 Supplements for All-Cause Mortality



Error bars indicate the 95% CI of the cumulative meta-analysis estimates as randomized patients accumulate through time. PUFAs indicates polyunsaturated fatty acids; RR, relative risk.

Additional Resources



For Systematic Review:

- University Library – Library Guide for Systematic Review https://libguides.library.usyd.edu.au/c.php?g=516278&p=3529739
- University Library – Workshops, Canvas modules and Consultations

For Meta-Analysis

- "Introduction to Meta-Analysis" Borenstein, Hedges, Higgins, Rothstein Link to library copy
- Cochrane.org has on-line learning, guides and handbooks http://training.cochrane.org/online-learning
- A guide to Prospective Meta-analysis (from USyd NHMRC CTC) https://www.bmj.com/content/367/bmj.l5342.short

Asking for more help

SIH Resources

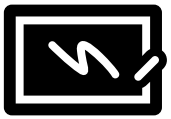
- Our website:

www.sydney.edu.au/research/facilities/sydney-informatics-hub.html
OR Google "Sydney Informatics Hub" with the "I'm feeling lucky" button

- SIH training
  - Sign up to our mailing list to be notified of upcoming training: mailman.sydney.edu.au/mailman/listinfo/computing_training
  - Hacky Hour
    www.sydney.edu.au/research/facilities/sydney-informatics-hub/workshops-and-training/hacky-hour.html
    OR Google "Sydney Hacky Hour"

- Other resources
  - OLE courses
  - Other University Accessible
    - Linkedin Learning [Meta-analysis for Data Science and Business Analytics]: https://linkedin.com/learning/

Acknowledging SIH



All University of Sydney resources are available to Sydney researchers **free of charge.** The use of the SIH services including the Artemis HPC and associated support and training warrants acknowledgement in any publications, conference proceedings or posters describing work facilitated by these services.

The continued acknowledgment of the use of SIH facilities ensures the sustainability of our services.

Suggested wording:

General acknowledgement:

"The authors acknowledge the technical assistance provided by the Sydney Informatics Hub, a Core Research Facility of the University of Sydney."

Acknowledging specific staff:

"The authors acknowledge the technical assistance of (name of staff) of the Sydney Informatics Hub, a Core Research Facility of the University of Sydney."

For further information about acknowledging the Sydney Informatics Hub, please contact us at **sih.info@sydney.edu.au**.

End of Workshop on Meta-analysis

- Thank you for your interest and attention
- Questions and comments welcome
- We appreciate your feedback via the on-line survey

Jim Matthews BEng MStat | Senior Consultant: Statistics
The University of Sydney
Sydney Informatics Hub | Core Research Facilities
Merewether Building (H04) | The University of Sydney | NSW | 2006
+61 412 246 271
Jim.Matthews@sydney.edu.au | sydney.edu.au