Cochrane Training

Introduction to meta-analysis

Trusted evidence.
Informed decisions.
Better health.

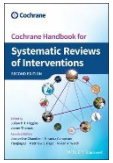
Cochrane Training

Steps of a Cochrane Review

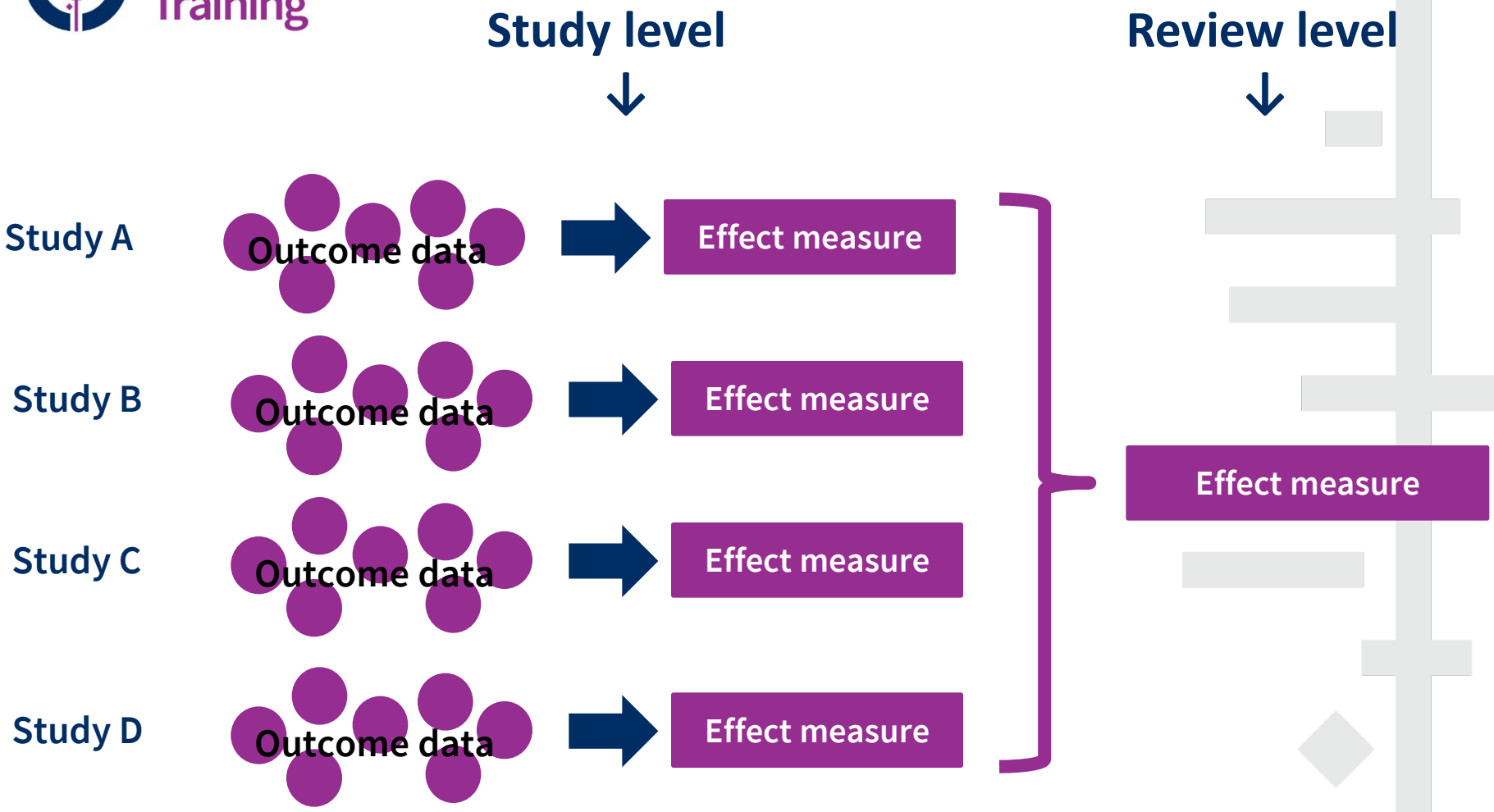
1. define the question
2. plan eligibility criteria
3. plan methods
4. search for studies
5. apply eligibility criteria
6. collect data
7. assess studies for risk of bias
8. **analyse and present results**
9. interpret results and draw conclusions
10. improve and update review

Session outline

- **principles of meta-analysis**
- steps in a meta-analysis
- presenting your results



See Chapter 10 of the Handbook

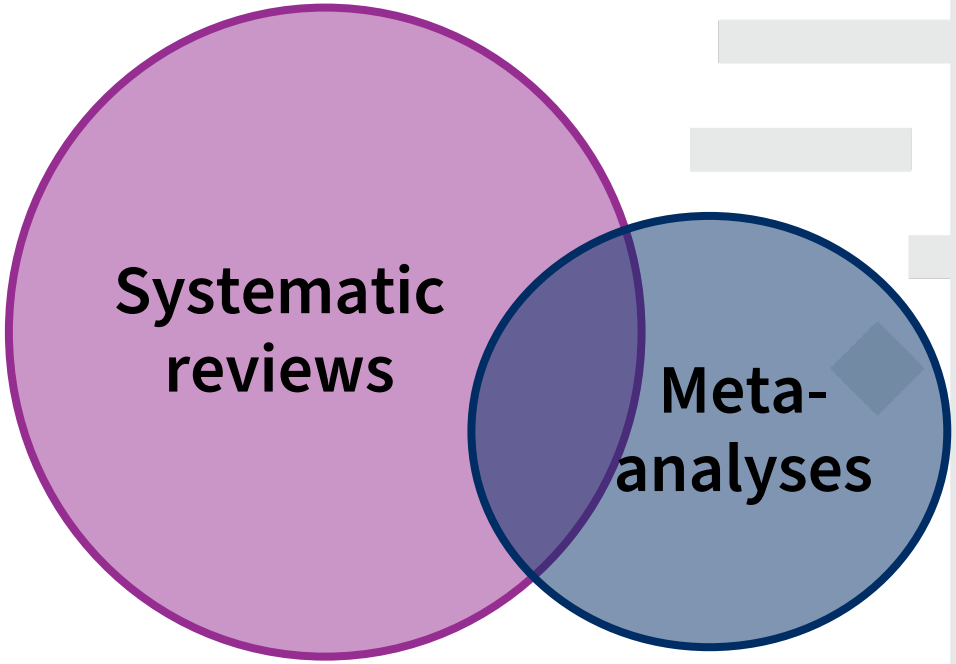


Source: Jo McKenzie & Miranda Cumpston

Cochrane Training

What is a meta-analysis?

- combines the results from two or more studies
- estimates an ‘average’ or ‘common’ effect
- optional part of a systematic review



Source: Julian Higgins

Why perform a meta-analysis?

- quantify treatment effects and their uncertainty
- increase power
- increase precision
- explore differences between studies
- settle controversies from conflicting studies
- generate new hypotheses

Source: Julian Higgins

When not to do a meta-analysis

- **mixing apples with oranges**
  - each included study must address the same question
    - consider comparison and outcomes
    - requires your subjective judgement
  - combining a broad mix of studies answers broad questions
  - answer may be meaningless and genuine effects may be obscured if studies are too diverse

Source: Julian Higgins

Cochrane Training

When not to do a meta-analysis

- **garbage in – garbage out**
  - a meta-analysis is only as good as the studies in it
  - if included studies are biased:
    - meta-analysis result will also be incorrect
    - will give more credibility and narrower confidence interval
  - if serious reporting biases present:
    - unrepresentative set of studies may give misleading result

Source: Julian Higgins

Cochrane Training

When can you do a meta-analysis?

- more than one study has measured an effect
- the studies are sufficiently similar to produce a meaningful and useful result
- the outcome has been measured in similar ways
- data are available in a format we can use

Cochrane Training

Session outline

- principles of meta-analysis
- **steps in a meta-analysis**
- presenting your results

Cochrane Training

Steps in a meta-analysis

- identify comparisons to be made
- identify outcomes to be reported and statistics to be used
- collect data from each relevant study
- combine the results to obtain the summary of effect
- explore differences between the studies
- interpret the results

Cochrane Training

Selecting comparisons

Hypothetical review: Caffeine for daytime drowsiness

caffeinated coffee vs decaffeinated coffee

- break your topic down into pair-wise comparisons
- each review may have one or many
- use your judgement to decide what to group together, and what should be a separate comparison

Selecting outcomes & effect measures

Hypothetical review: Caffeine for daytime drowsiness

caffeinated coffee vs decaffeinated coffee

- **asleep at end of trial (RR)**
- **irritability (MD/SMD)**
- **headaches (RR)**

- for each comparison, select outcomes
- for each outcome, select an effect measure
  - may depend on the available data from included studies

Calculating the summary result

- collect a summary statistic from each contributing study
- how do we bring them together?
  - treat as one big study – add intervention & control data?
    - breaks randomisation, will give the wrong answer
  - simple average?
    - weights all studies equally – some studies closer to the truth
  - weighted average

Weighting studies

- more weight to the studies which give more information
  - more participants, more events, narrower confidence interval
  - Base weights on study uncertainty or standard error

Larger trials → Smaller standard error

- BUT not always the case

Smaller standard error → Larger weight

SE small → 1 ÷ SE large

Weighting studies

- more weight to the studies which give more information
  - calculated using the effect estimate and its variance

- Could use 1 ÷ SE as the weight
- Actually use the inverse-variance method:

$$\text{weight} = \frac{1}{\text{variance of estimate}} = \frac{1}{SE^2}$$

$$\text{pooled estimate} = \frac{\text{sum of (estimate} \times \text{weight)}}{\text{sum of weights}}$$

Cochrane Training

For example

Headache	Caffeine	Decaf	Weight
Amore-Coffea 2000	2/31	10/34	
Deliciozza 2004	10/40	9/40	
Mama-Kaffa 1999	12/53	9/61	
Morrocona 1998	3/15	1/17	
Norscafe 1998	19/68	9/64	
Oohlahlazza 1998	4/35	2/37	
Piazza-Allerta 2003	8/35	6/37	

Cochrane Training

For example

Headache	Caffeine	Decaf	Weight
Amore-Coffea 2000	2/31	10/34	6.6%
Deliciozza 2004	10/40	9/40	21.9%
Mama-Kaffa 1999	12/53	9/61	22.2%
Morrocona 1998	3/15	1/17	2.9%
Norscafe 1998	19/68	9/64	26.4%
Oohlahlazza 1998	4/35	2/37	5.1%
Piazza-Allerta 2003	8/35	6/37	14.9%

Cochrane Training

Meta-analysis options

- for dichotomous or continuous data
  - inverse-variance
    - straightforward, general method
- for dichotomous data only
  - Mantel-Haenszel (default)
    - good with few events – common in Cochrane reviews
    - weighting system depends on effect measure
  - Peto
    - for odds ratios only
    - good with few events and small effect sizes (OR close to 1)

Cochrane Training

Meta-analysis options

Name	Headache
Data source	**Custom input**
Data type	**Dichotomous**
Intervention group 1	Caffeinated coffee
Intervention group 2	Decaffeinated coffee

Statistical settings

Statistical method	Inverse variance
Effect measure	Risk ratio
Analysis model	Fixed effect
Totals	Totals and subtotals

☑ Test for subgroup differences

☐ Swap event and non-event

Study confidence interval	95%
Total confidence interval	95%

Cochrane Training

Session outline

- principles of meta-analysis
- steps in a meta-analysis
- **presenting your results**

A forest of lines

Forest plots

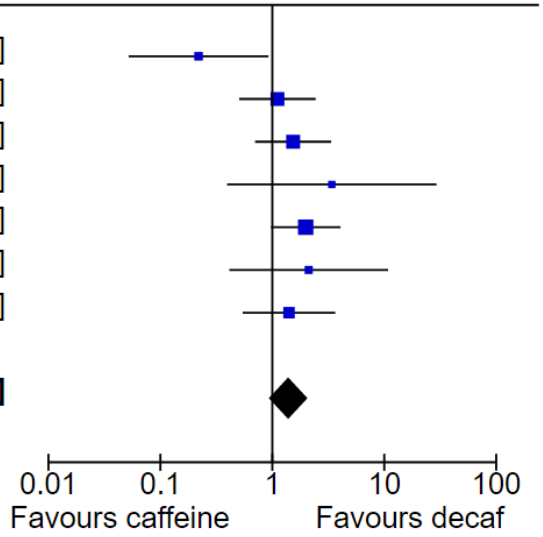
Headache at 24 hours

Study or Subgroup	Caffeinated coffee Events	Caffeinated coffee Total	Decaffeinated coffee Events	Decaffeinated coffee Total	Weight	Risk ratio IV, Fixed, 95% CI
Amore-Coffea 2000	2	31	10	34	6.6%	0.22 [0.05 , 0.92]
Deliciozza 2004	10	40	9	40	21.9%	1.11 [0.51 , 2.44]
Mama-Kaffa 1999	12	53	9	61	22.2%	1.53 [0.70 , 3.35]
Morrocona 1998	3	15	1	17	2.9%	3.40 [0.39 , 29.31]
Norscafe 1998	19	68	9	64	26.4%	1.99 [0.97 , 4.07]
Oohlahlazza 1998	4	35	2	37	5.1%	2.11 [0.41 , 10.83]
Piazza-Allerta 2003	8	35	6	37	14.9%	1.41 [0.54 , 3.65]
Total (95% CI)		**277**		**290**	**100.0%**	**1.38 [0.96 , 2.00]**
Total events:	58		46			

Heterogeneity: Chi² = 8.58, df = 6 (P = 0.20); I² = 30%

Test for overall effect: Z = 1.73 (P = 0.08)

Test for subgroup differences: Not applicable



- headings explain the comparison

Forest plots

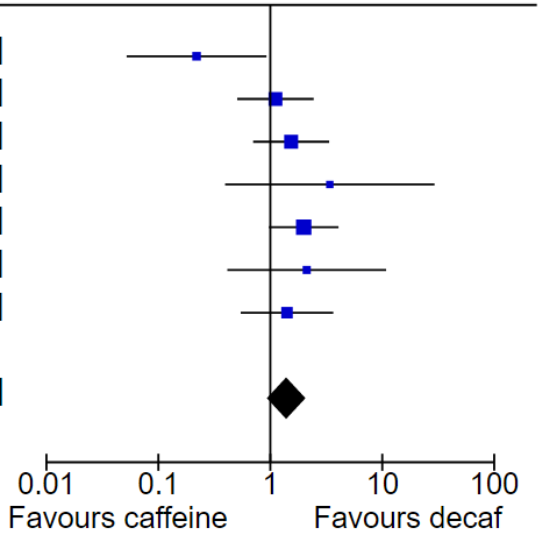
Headache at 24 hours

Study or Subgroup	Caffeinated coffee Events	Total	Decaffeinated coffee Events	Total	Weight	Risk ratio IV, Fixed, 95% CI
Amore-Coffea 2000	2	31	10	34	6.6%	0.22 [0.05 , 0.92]
Deliciozza 2004	10	40	9	40	21.9%	1.11 [0.51 , 2.44]
Mama-Kaffa 1999	12	53	9	61	22.2%	1.53 [0.70 , 3.35]
Morrocona 1998	3	15	1	17	2.9%	3.40 [0.39 , 29.31]
Norscafe 1998	19	68	9	64	26.4%	1.99 [0.97 , 4.07]
Oohlahlazza 1998	4	35	2	37	5.1%	2.11 [0.41 , 10.83]
Piazza-Allerta 2003	8	35	6	37	14.9%	1.41 [0.54 , 3.65]
Total (95% CI)		**277**		**290**	**100.0%**	**1.38 [0.96 , 2.00]**
Total events:	58		46			

Heterogeneity: Chi² = 8.58, df = 6 (P = 0.20); I² = 30%

Test for overall effect: Z = 1.73 (P = 0.08)

Test for subgroup differences: Not applicable



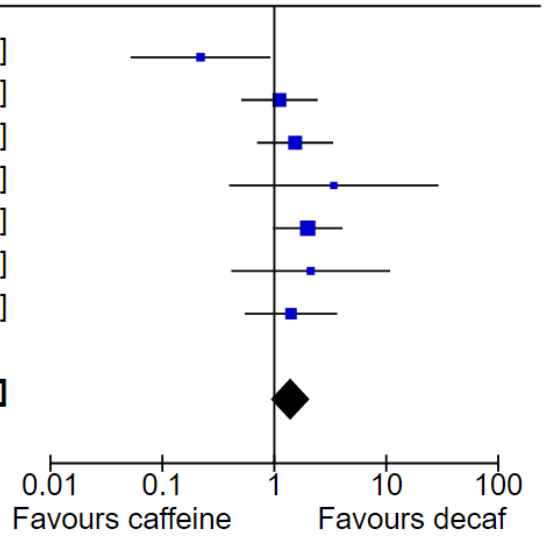
- list of included studies

Forest plots

Headache at 24 hours

Study or Subgroup	Caffeinated coffee Events	Total	Decaffeinated coffee Events	Total	Weight	Risk ratio IV, Fixed, 95% CI
Amore-Coffea 2000	2	31	10	34	6.6%	0.22 [0.05 , 0.92]
Deliciozza 2004	10	40	9	40	21.9%	1.11 [0.51 , 2.44]
Mama-Kaffa 1999	12	53	9	61	22.2%	1.53 [0.70 , 3.35]
Morrocona 1998	3	15	1	17	2.9%	3.40 [0.39 , 29.31]
Norscafe 1998	19	68	9	64	26.4%	1.99 [0.97 , 4.07]
Oohlahlazza 1998	4	35	2	37	5.1%	2.11 [0.41 , 10.83]
Piazza-Allerta 2003	8	35	6	37	14.9%	1.41 [0.54 , 3.65]
Total (95% CI)		**277**		**290**	**100.0%**	**1.38 [0.96 , 2.00]**
Total events:	58		46			

Heterogeneity: Chi² = 8.58, df = 6 (P = 0.20); I² = 30%
Test for overall effect: Z = 1.73 (P = 0.08)
Test for subgroup differences: Not applicable



- raw data for each study

Forest plots

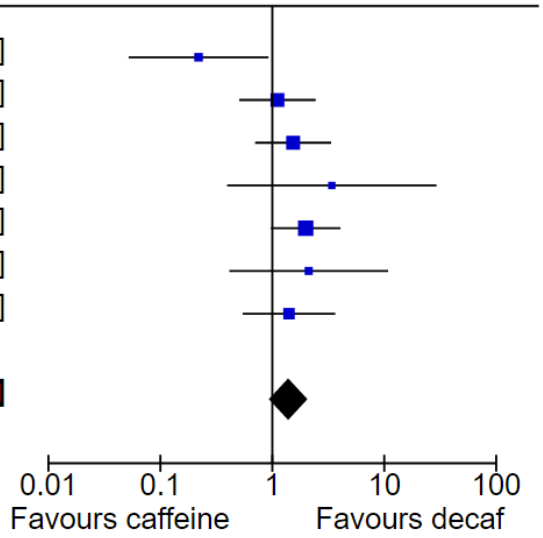
Headache at 24 hours

Study or Subgroup	Caffeinated coffee Events	Total	Decaffeinated coffee Events	Total	Weight	Risk ratio IV, Fixed, 95% CI
Amore-Coffea 2000	2	31	10	34	6.6%	0.22 [0.05 , 0.92]
Deliciozza 2004	10	40	9	40	21.9%	1.11 [0.51 , 2.44]
Mama-Kaffa 1999	12	53	9	61	22.2%	1.53 [0.70 , 3.35]
Morrocona 1998	3	15	1	17	2.9%	3.40 [0.39 , 29.31]
Norscafe 1998	19	68	9	64	26.4%	1.99 [0.97 , 4.07]
Oohlahlazza 1998	4	35	2	37	5.1%	2.11 [0.41 , 10.83]
Piazza-Allerta 2003	8	35	6	37	14.9%	1.41 [0.54 , 3.65]
Total (95% CI)		**277**		**290**	**100.0%**	**1.38 [0.96 , 2.00]**
Total events:	58		46			

Heterogeneity: Chi² = 8.58, df = 6 (P = 0.20); I² = 30%

Test for overall effect: Z = 1.73 (P = 0.08)

Test for subgroup differences: Not applicable



- total data for all studies

Forest plots

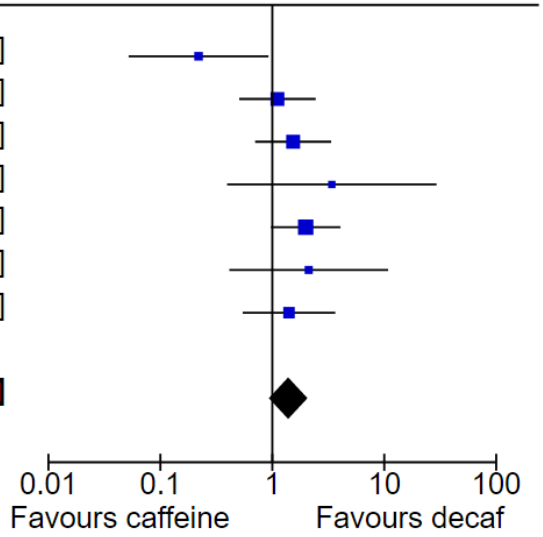
Headache at 24 hours

Study or Subgroup	Caffeinated coffee Events	Total	Decaffeinated coffee Events	Total	Weight	Risk ratio IV, Fixed, 95% CI
Amore-Coffea 2000	2	31	10	34	6.6%	0.22 [0.05 , 0.92]
Deliciozza 2004	10	40	9	40	21.9%	1.11 [0.51 , 2.44]
Mama-Kaffa 1999	12	53	9	61	22.2%	1.53 [0.70 , 3.35]
Morrocona 1998	3	15	1	17	2.9%	3.40 [0.39 , 29.31]
Norscafe 1998	19	68	9	64	26.4%	1.99 [0.97 , 4.07]
Oohlahlazza 1998	4	35	2	37	5.1%	2.11 [0.41 , 10.83]
Piazza-Allerta 2003	8	35	6	37	14.9%	1.41 [0.54 , 3.65]
Total (95% CI)		**277**		**290**	**100.0%**	**1.38 [0.96 , 2.00]**
Total events:	58		46			

Heterogeneity: Chi² = 8.58, df = 6 (P = 0.20); I² = 30%

Test for overall effect: Z = 1.73 (P = 0.08)

Test for subgroup differences: Not applicable



- weight given to each study

Forest plots

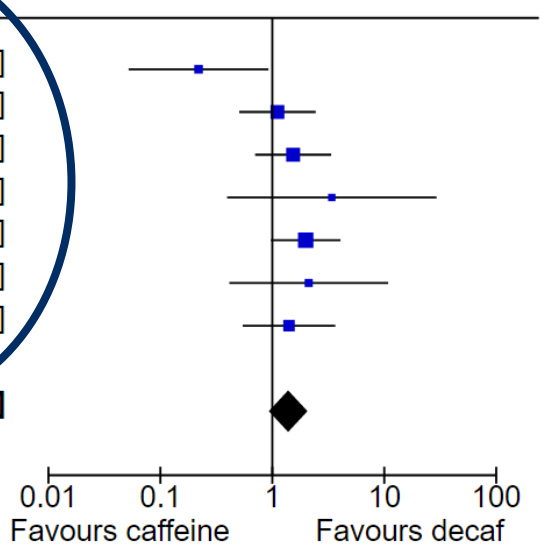
Headache at 24 hours

Study or Subgroup	Caffeinated coffee Events	Total	Decaffeinated coffee Events	Total	Weight	Risk ratio IV, Fixed, 95% CI
Amore-Coffea 2000	2	31	10	34	6.6%	0.22 [0.05 , 0.92]
Deliciozza 2004	10	40	9	40	21.9%	1.11 [0.51 , 2.44]
Mama-Kaffa 1999	12	53	9	61	22.2%	1.53 [0.70 , 3.35]
Morrocona 1998	3	15	1	17	2.9%	3.40 [0.39 , 29.31]
Norscafe 1998	19	68	9	64	26.4%	1.99 [0.97 , 4.07]
Oohlahlazza 1998	4	35	2	37	5.1%	2.11 [0.41 , 10.83]
Piazza-Allerta 2003	8	35	6	37	14.9%	1.41 [0.54 , 3.65]
Total (95% CI)		**277**		**290**	**100.0%**	**1.38 [0.96 , 2.00]**
Total events:	58		46			

Heterogeneity: Chi² = 8.58, df = 6 (P = 0.20); I² = 30%

Test for overall effect: Z = 1.73 (P = 0.08)

Test for subgroup differences: Not applicable



- effect estimate for each study, with CI

Forest plots

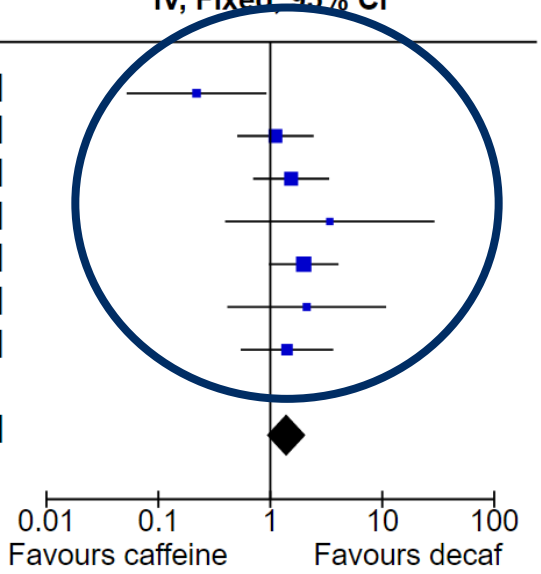
Headache at 24 hours

Study or Subgroup	Caffeinated coffee Events	Total	Decaffeinated coffee Events	Total	Weight	Risk ratio IV, Fixed, 95% CI
Amore-Coffea 2000	2	31	10	34	6.6%	0.22 [0.05 , 0.92]
Deliciozza 2004	10	40	9	40	21.9%	1.11 [0.51 , 2.44]
Mama-Kaffa 1999	12	53	9	61	22.2%	1.53 [0.70 , 3.35]
Morrocona 1998	3	15	1	17	2.9%	3.40 [0.39 , 29.31]
Norscafe 1998	19	68	9	64	26.4%	1.99 [0.97 , 4.07]
Oohlahlazza 1998	4	35	2	37	5.1%	2.11 [0.41 , 10.83]
Piazza-Allerta 2003	8	35	6	37	14.9%	1.41 [0.54 , 3.65]
Total (95% CI)		**277**		**290**	**100.0%**	**1.38 [0.96 , 2.00]**
Total events:	58		46			

Heterogeneity: Chi² = 8.58, df = 6 (P = 0.20); I² = 30%

Test for overall effect: Z = 1.73 (P = 0.08)

Test for subgroup differences: Not applicable



- effect estimate for each study, with CI

Forest plots

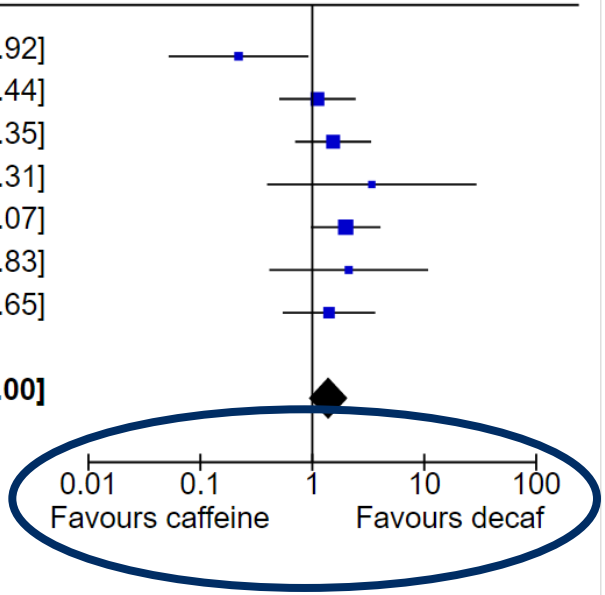
Headache at 24 hours

Study or Subgroup	Caffeinated coffee Events	Total	Decaffeinated coffee Events	Total	Weight	Risk ratio IV, Fixed, 95% CI
Amore-Coffea 2000	2	31	10	34	6.6%	0.22 [0.05 , 0.92]
Deliciozza 2004	10	40	9	40	21.9%	1.11 [0.51 , 2.44]
Mama-Kaffa 1999	12	53	9	61	22.2%	1.53 [0.70 , 3.35]
Morrocona 1998	3	15	1	17	2.9%	3.40 [0.39 , 29.31]
Norscafe 1998	19	68	9	64	26.4%	1.99 [0.97 , 4.07]
Oohlahlazza 1998	4	35	2	37	5.1%	2.11 [0.41 , 10.83]
Piazza-Allerta 2003	8	35	6	37	14.9%	1.41 [0.54 , 3.65]
Total (95% CI)		**277**		**290**	**100.0%**	**1.38 [0.96 , 2.00]**
Total events:	58		46			

Heterogeneity: Chi² = 8.58, df = 6 (P = 0.20); I² = 30%

Test for overall effect: Z = 1.73 (P = 0.08)

Test for subgroup differences: Not applicable



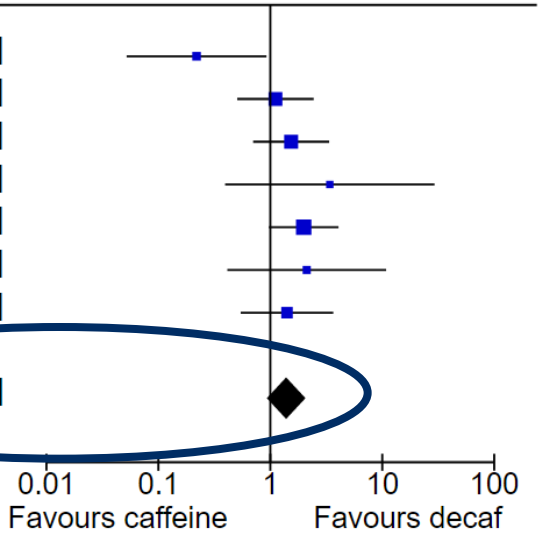
- scale and direction of benefit

Forest plots

Headache at 24 hours

Study or Subgroup	Caffeinated coffee Events	Total	Decaffeinated coffee Events	Total	Weight	Risk ratio IV, Fixed, 95% CI
Amore-Coffea 2000	2	31	10	34	6.6%	0.22 [0.05 , 0.92]
Deliciozza 2004	10	40	9	40	21.9%	1.11 [0.51 , 2.44]
Mama-Kaffa 1999	12	53	9	61	22.2%	1.53 [0.70 , 3.35]
Morrocona 1998	3	15	1	17	2.9%	3.40 [0.39 , 29.31]
Norscafe 1998	19	68	9	64	26.4%	1.99 [0.97 , 4.07]
Oohlahlazza 1998	4	35	2	37	5.1%	2.11 [0.41 , 10.83]
Piazza-Allerta 2003	8	35	6	37	14.9%	1.41 [0.54 , 3.65]
Total (95% CI)		**277**		**290**	**100.0%**	**1.38 [0.96 , 2.00]**
Total events:	58		46			

Heterogeneity: Chi² = 8.58, df = 6 (P = 0.20); I² = 30%
Test for overall effect: Z = 1.73 (P = 0.08)
Test for subgroup differences: Not applicable



- pooled effect estimate for all studies, with CI

Cochrane Training

Confidence intervals

- Always present estimate with a confidence interval (CI)
- Effect estimate from study sample – best guess of effect
- Estimating effect in greater population - uncertainty
- 95% CIs are calculated to describe uncertainty around the effect estimate
- Describes a range of values we can be 95% certain includes the true effect
- Precision
  - Narrow CI (e.g. 0.70 to 0.80) indicates that the effect size is known precisely
  - Wide CI (e.g. 0.60 to 0.93) indicates greater uncertainty
  - Very wide CI (e.g. 0.50 to 1.10) indicates that we have very little knowledge about the effect

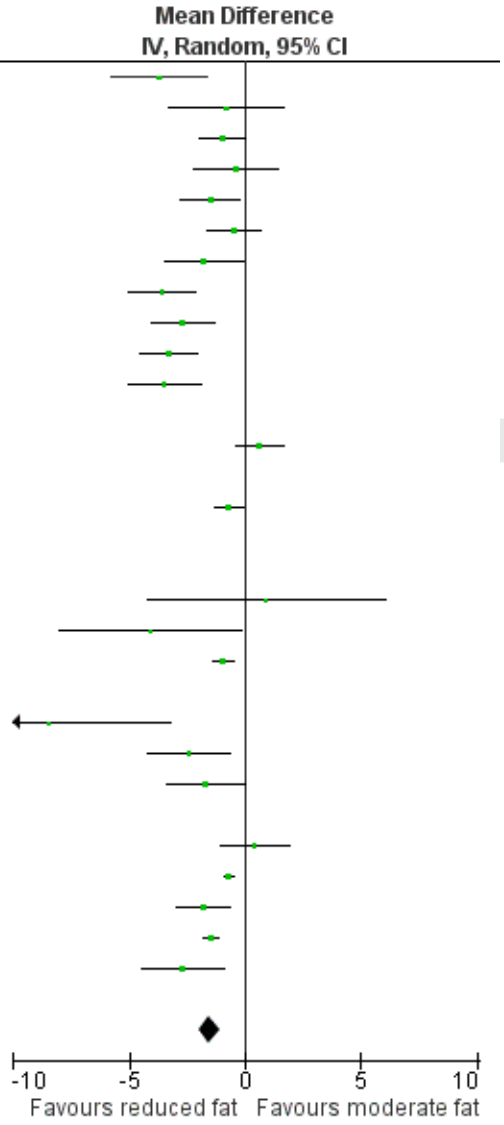
Confidence intervals

- Largely affected by sample size
  - Larger sample = narrower 95%CIs in individual studies
  - In meta-analysis, more studies will lead to a narrower CI if results are similar
  - If results are conflicting, more studies may lead to a wider CI
- Significance
  - if the CI includes the null value
    - rarely means evidence of no effect
    - effect cannot be confirmed or refuted by the available evidence
  - consider what level of change is clinically important

Considering clinical significance

Study or Subgroup	Reduced fat Mean	SD	Total	Usual or modified fat Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Auckland reduced fat 1999	-1.6	5.4	48	2.13	5	51	2.8%	-3.73 [-5.78, -1.68]
BDIT Pilot Studies 1996	59.6	7.3	76	60.4	8.4	78	2.1%	-0.80 [-3.28, 1.68]
Bloemberg 1991	-0.94	2.68	39	0.06	1.86	40	5.5%	-1.00 [-2.02, 0.02]
BRIDGES 2001	0.1	4.85	48	0.5	4.07	46	3.3%	-0.40 [-2.21, 1.41]
Canadian DBCP 1997	62	9.1	388	63.5	9.4	401	4.6%	-1.50 [-2.79, -0.21]
de Bont 1981 non-obese	-0.4	2.8	36	0.1	2	29	5.0%	-0.50 [-1.67, 0.67]
de Bont 1981 obese	-2.7	3.6	34	-0.9	3.5	35	3.5%	-1.80 [-3.48, -0.12]
DEER 1998 exercise men	-4.2	4.2	48	-0.6	3.1	47	4.0%	-3.60 [-5.08, -2.12]
DEER 1998 exercise women	-3.1	3.7	43	-0.4	2.5	43	4.5%	-2.70 [-4.03, -1.37]
DEER 1998 no exercise men	-2.8	3.5	49	0.5	2.7	46	4.7%	-3.30 [-4.55, -2.05]
DEER 1998 no exercise wom	-2.7	3.5	46	0.8	4.2	45	3.8%	-3.50 [-5.09, -1.91]
Diet and Hormone Study 2003	-0.68	0	81	-0.14	0	96		Not estimable
Kentucky Low Fat 1990	1.06	2.49	47	0.44	2.68	51	5.4%	0.62 [-0.40, 1.64]
MeDiet 2006	-1.3	0	51	-0.6	0	55		Not estimable
MSFAT 1995	0.4	2.36	117	1.12	2.36	103	6.8%	-0.72 [-1.34, -0.10]
NDHS Open 1st L&M 1968	-2.45	0	332	-1.91	0	348		Not estimable
NDHS Open 2nd L&M 1968	-1.8	0	179	-1.2	0	215		Not estimable
Nutrition & Breast Health	67.3	13.8	47	66.4	12	50	0.6%	0.90 [-4.26, 6.06]
Pilkington 1960	66.7	5.9	12	70.8	5.2	23	1.0%	-4.10 [-8.06, -0.14]
Polyp Prevention 1996	-0.65	5.22	943	0.31	5.22	943	7.3%	-0.96 [-1.43, -0.49]
Rivellese 1994	-1.8	0	27	-1.6	0	17		Not estimable
Simon Low Fat Breast CA	63.4	11.1	34	71.9	11.7	38	0.6%	-8.50 [-13.77, -3.23]
Strychar 2009	-0.83	3	15	1.6	1.8	15	3.3%	-2.43 [-4.20, -0.66]
Swedish Breast CA 1990	-0.4	5.5	63	1.3	5.5	106	3.5%	-1.70 [-3.41, 0.01]
Veterans Dermatology 1994	-2	0	38	0.5	0	58		Not estimable
WHEL 2007	74.1	19.53	1308	73.7	19.2	1313	4.0%	0.40 [-1.08, 1.88]
WHI 2006	-0.8	10.1	16297	-0.1	10.1	25056	7.9%	-0.70 [-0.90, -0.50]
WHT Feasibility 1990	-1.91	4.9	159	-0.08	4.3	102	5.1%	-1.83 [-2.96, -0.70]
WHT:FSMP 2003	-1.8	4	1325	-0.3	4.2	883	7.6%	-1.50 [-1.85, -1.15]
WINS 1993	-2.7	15.3	386	0	15.3	998	3.3%	-2.70 [-4.50, -0.90]
Total (95% CI)			**22316**			**31331**	**100.0%**	**-1.54 [-1.97, -1.12]**

Heterogeneity: Tau² = 0.58; Chi² = 99.49, df = 23 (P < 0.00001); I² = 77%
Test for overall effect: Z = 7.14 (P < 0.00001)



Hooper L, Abdelhamid A, Bunn D, Brown T, Summerbell CD, Skeaff CM. Effects of total fat intake on body weight. Cochrane Database of Systematic Reviews 2015, Issue 8. Art. No.: CD011834. DOI: 10.1002/14651858.CD011834

Cochrane Training

Presenting data in your review

- present outcomes in consistent order throughout
  - Abstract, Methods, Results, data
- forest plots
  - key forest plots linked as figures
    - usually primary outcomes
  - all forest plots will be published as supplementary data
  - avoid forest plots with only one study
- may also add other data tables
  - results of single studies
    - summary data for each group, effect estimates, confidence intervals
  - non-standard data

Cochrane Training

What to include in the protocol

- how will you decide whether a meta-analysis is appropriate?
- meta-analysis model to be used

Cochrane Training

Take home message

- there are several advantages to performing a meta-analysis but it is not always possible (or appropriate)
- plan your analysis carefully, including comparisons, outcomes and meta-analysis methods
- forest plots display the results of meta-analyses graphically
- interpret your results with caution

References

- Deeks JJ, Higgins JPT, Altman DG (editors). **Chapter 10: Analysing data and undertaking meta-analyses.** In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.S
- Schünemann HJ, Higgins JPT, Vist GE, Glasziou P, Akl EA, Skoetz N, Guyatt GH. **Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence**. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.

Acknowledgements

- **Compiled by Miranda Cumpston**
- **Based on materials by Sally Hopewell, Julian Higgins, the Cochrane Statistical Methods Group and Cochrane Netherlands**
- **Approved by the Convenors of Cochrane Methods Groups**