NIPH
Norwegian Institute of Public Health

# Analyzing data and conducting a meta-analysis

Eva Denison
Senior researcher, NIPH

1

## Steps in conducting a SR

1.Formulate the question

2.Define criteria for inclusion- and exclusion

3.Identify (locate) studies

4.Select studies

5.Assess methodological quality of studies (bias)

6.Extract data

**7.Analyse data**

8.GRADE

9.Present and interpret results



2

NIPH
Norwegian Institute of Public Health

Approaches to data synthesis in systematic reviews

- Narrative
- Quantitative


3

# Sources

Cochrane Training
Trusted evidence.
Informed decisions.
Better health.

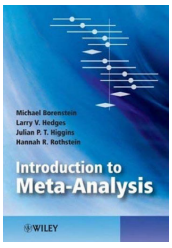
Search...

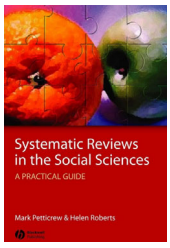
Online learning | Learning events | Guides and handbooks | Trainers' Network | Log in

Chapter 10: Analysing data and undertaking meta-analyses

Search Handbook

Jonathan J Deeks, Julian PT Higgins, Douglas G Altman; on behalf of the Cochrane Statistical Methods Group





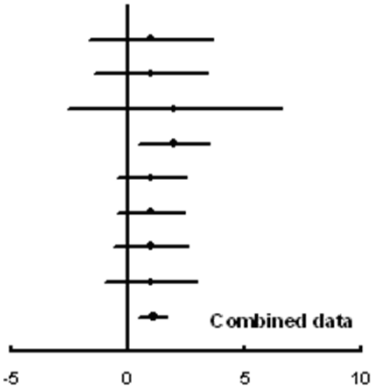

4

# Two approaches

## Narrative

**Results** 22 studies met the inclusion criteria. We found some evidence that targeted behaviour change programmes can change the behaviour of motivated subgroups, resulting (in the largest study) in a shift of around 5% of all trips at a population level. Single studies of commuter subsidies and a new railway station also showed positive effects. The balance of best available evidence about publicity campaigns, engineering measures, and other interventions suggests that they have not been effective. Participants in trials of active commuting experienced short term improvements in certain measures of health and fitness, but we found no good evidence on effects on health of any effective intervention at population level. **Conclusions** The best available evidence of effectiveness in promoting a modal shift is for targeted behaviour change programmes, but the social distribution of their effects is unclear and some other types of intervention have yet to be rigorously evaluated.

## Quantitative – meta-analysis




In between: Forest plot without combining data



5

# Two approaches

## Narrative

- When quantitative analysis is not deemed appropriate
  - «Clinical heterogeneity», e.g.
    - Complex interventions
    - Different settings
    - Different measurement methods and length of follow-up
- Studies with qualitative data

## Quantitative

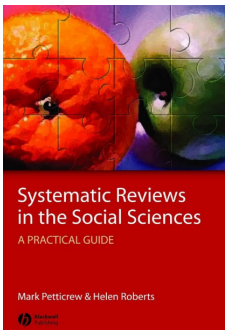
- Statistical synthesis of quantitative data
  - Improved power
    - Increased N due to pooling of studies
  - Improved precision
    - Narrower confidence intervals around the population estimate
  - Beware
    - Methodological bias, heterogeneity, publication bias



6

# Narrative synthesis in three steps

1. Organize the description of the studies in logical categories
2. Analyse the findings within each category
3. Summarize the findings across all categories






7

# Narrative synthesis

- All data are summarized in tables
  - PIO/PICO/PEO/PECO
  - Methodological quality
  - Findings
  - Context
  - Other information of interest
- The tables themselves are not the synthesis but the basis for the synthesis



8

# «Vote counting»

- How many studies have positive or negative results?
- NOT recommended – does not account for
  - The size of the sample
  - Study methods and quality
  - Qualitative differences between the studies
  - Interactions between the variables in the studies



9

Cite this article as: BMJ, doi:10.1136/bmj.38216.714560.55 (published 22 September 2004)

Papers

## Promoting walking and cycling as an alternative to using cars: systematic review

David Ogilvie, Matt Egan, Val Hamilton, Mark Petticrew



10

**Validity score**
10 items
- Selection
- Sample size
- Attrition rate
- Reporting
- Performance
- Follow-up

Summary of evidence of effectiveness of interventions to promote modal shift

| Study | Validity score | Nature of comparison | Evidence for shift from cars towards walking and cycling*: Significant positive effect | Positive effect of uncertain significance | Inconclusive or no effect | Negative effect of uncertain significance |
|---|---|---|---|---|---|---|
| **Targeted behaviour change programmes** | | | | | | |
| Glasgow[w1 w2] | 9 | Controlled | Yes | — | — | — |
| Perth, Australia (TravelSmart)[w3-9] | 7 | Controlled | Yes | — | — | — |
| Frome (TravelSmart pilot)[w10] | 9 | Controlled | — | Yes | — | — |
| Gloucester (TravelSmart pilot)[w11] | 9 | Controlled | — | Yes | — | — |
| Århus[w12-14] | 7 | Uncontrolled | — | Yes | — | — |
| Adelaide[w15-18] | 4 | Uncontrolled | — | — | Yes | — |
| **Publicity campaigns and agents of change** | | | | | | |
| Camden-Islington[w19] | 8 | Controlled | — | — | Yes | — |
| Maidstone[w20] | 7 | Controlled | — | — | Yes | — |
| Phoenix[w21] | 5 | Uncontrolled | — | Yes | — | — |
| Eugene[w22] | 4 | Uncontrolled | — | — | Yes | — |
| **Engineering measures** | | | | | | |
| Delft[w23-30] | 7 | Controlled | — | Yes | — | — |
| Detmold-Rosenheim[w31-33] | 6 | Uncontrolled | — | — | — | Yes |
| Stockton[w34] | 5 | Uncontrolled | — | — | — | Yes |
| England (20 mph (30 km/h) zones)[w35] | 5 | Uncontrolled | — | — | Yes | — |
| Boston[w38-40] | 4 | Uncontrolled | — | Yes | — | — |
| England (bypasses)[w36 w37] | 3 | Uncontrolled | — | — | — | Yes |
| **Financial incentives** | | | | | | |
| California (cashing out)[w41 w42] | 8 | Controlled | Yes | — | — | — |
| Trondheim[w43 w44] | 7 | Uncontrolled | — | — | — | Yes |
| **Providing alternative services** | | | | | | |
| San Francisco[w45-47] | 7 | Controlled | — | — | Yes | — |
| Voorhout[w48] | 7 | Uncontrolled | Yes | — | — | — |
| California (telecommuting)[w49] | 4 | Controlled | — | — | — | Yes |

*No studies had significant negative effects.



11

# Exercise 1 in break-out groups

## Narrative synthesis



12

# Synthesis

**Results** 22 studies met the inclusion criteria. We found some evidence that targeted behaviour change programmes can change the behaviour of motivated subgroups, resulting (in the largest study) in a shift of around 5% of all trips at a population level. Single studies of commuter subsidies and a new railway station also showed positive effects. The balance of best available evidence about publicity campaigns, engineering measures, and other interventions suggests that they have not been effective. Participants in trials of active commuting experienced short term improvements in certain measures of health and fitness, but we found no good evidence on effects on health of any effective intervention at population level. **Conclusions** The best available evidence of effectiveness in promoting a modal shift is for targeted behaviour change programmes, but the social distribution of their effects is unclear and some other types of intervention have yet to be rigorously evaluated.



13



14

# Meta-analysis

- The statistical combination of results from two or more studies
  - Potential advantages of meta-analyses
    - increase in power
    - improvement in precision
    - ability to answer questions not posed by individual studies
    - opportunity to settle controversies arising from conflicting claims
  - Potential to mislead seriously
    - specific study designs
    - within-study biases
    - variation across studies
    - reporting biases



15

# Steps in performing a meta-analysis

- Define a clear and focused topic for the review
- Establish inclusion- and exclusion criteria
- Locate all studies relevant to the topic
- Abstract information from the publications
- Assess risk of bias
- Carry out a descriptive analysis
- Carry out a statistical analysis
- Interprete the results



16

# Effect estimates

- Dichotomous outcomes
  - Relative risk (RR)
  - Odds ratio (OR)
  - Hazard ratio (HR)
- Continuous outcomes
  - Mean difference (MD)
  - Standardized mean difference (SMD)
    - Only shows direction and magnitude of effects
    - 0.2 small effect; 0.5 medium effect; 0.8 large effect



17

# Also consider

If you have another core question than effect

- Measures of incidence
- Measures of disease risk
- Measures of association
- Measures of impact

- …and consider consulting a statistician!



18

# Statistical models

- Fixed effect – we assume that
  - All studies estimate the same intervention effect
  - All variation in observed effects are due to sampling error
- Random effect – we assume that
  - Intervention effects may vary across studies, e.g. due to different mix of participants and implementation of interventions
  - Distribution of effects across studies



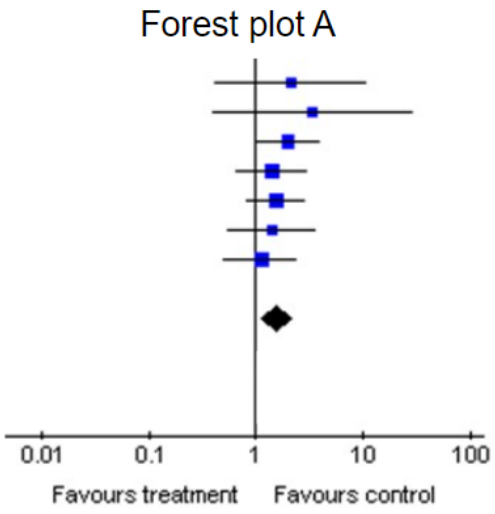
19

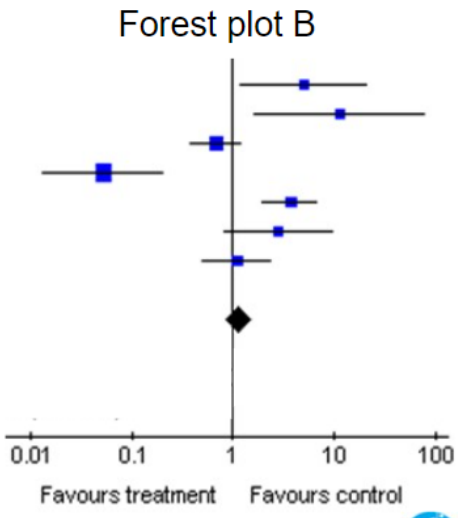
# Heterogeneity

- «Clinical»
  - Comparing apples and oranges
  - PICO, broad inclusion criteria
  - Splitting/lumping
  - Is it appropriate to conduct meta-analysis?
- Statistical – what proportion of the variation that cannot be explained by random variation
  - Calculated in the meta-analysis
  - **I-square, Chi-square ($p < 0.10$ indicates statistical significance)**
- Statistical - the extent of variation among the effects observed in different studies
  - Calculated in the meta-analysis
  - **Tau square**



20




cochrane training



21

# Interpretation of heterogeneity

| $I^2$ | Interpretation |
|---|---|
| 0 – 40 % | Might not be important |
| 30 – 60 % | May represent moderate heterogeneity |
| 50 – 90 % | May represent substantial heterogeneity |
| 75 – 100 % | Considerable heterogeneity |

Importance

Size and direction of observed effects

Strength of evidence (p-value from Chi-square test)



22

# Strategies for dealing with heterogeneity

| Strategy | Possible solution |
|---|---|
| Check data | Correct extraction errors or choice of unit |
| Evade | Try other effect measures |
| Ignore | Don't! |
| Give up | Drop meta-analysis |
| Explore | Does the heterogeneity disappear in subgroup- and sensitivity analyses? |
| Embrace | Use a statistical model that opens for variation between primary studies |



23

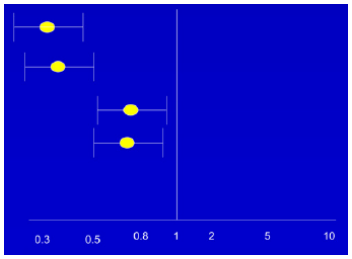
# Exercise 2 in break-out groups

## Heterogeneity and statistical models

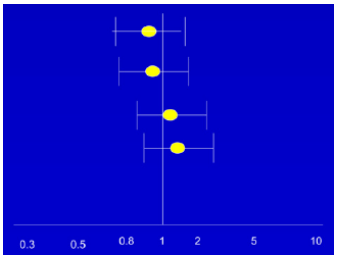


24

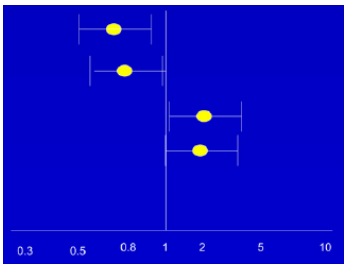
## Is it statistical heterogeneity?




No overlap,
same direction of effect




Overlap,
different directions of effect




No overlap,
different directions of effect



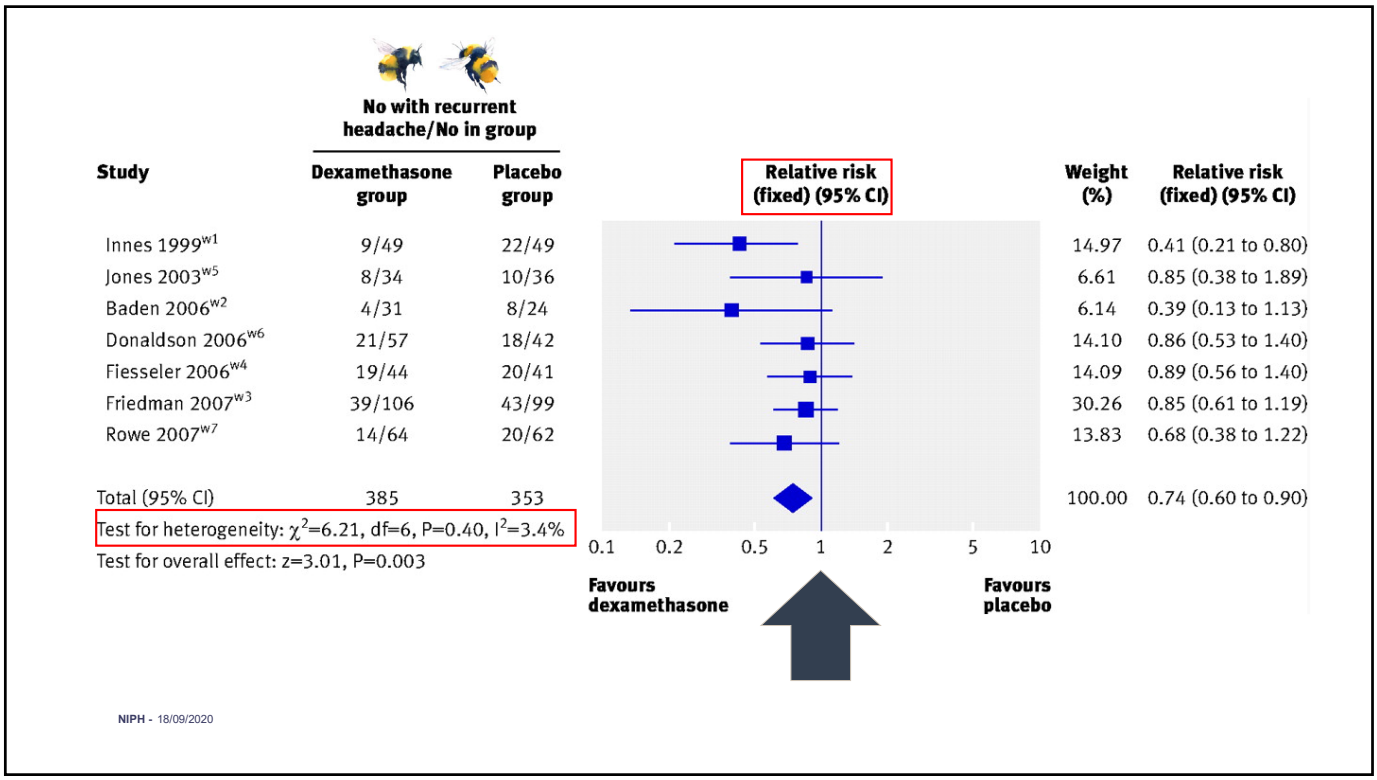
25

# Analyses and forest plots

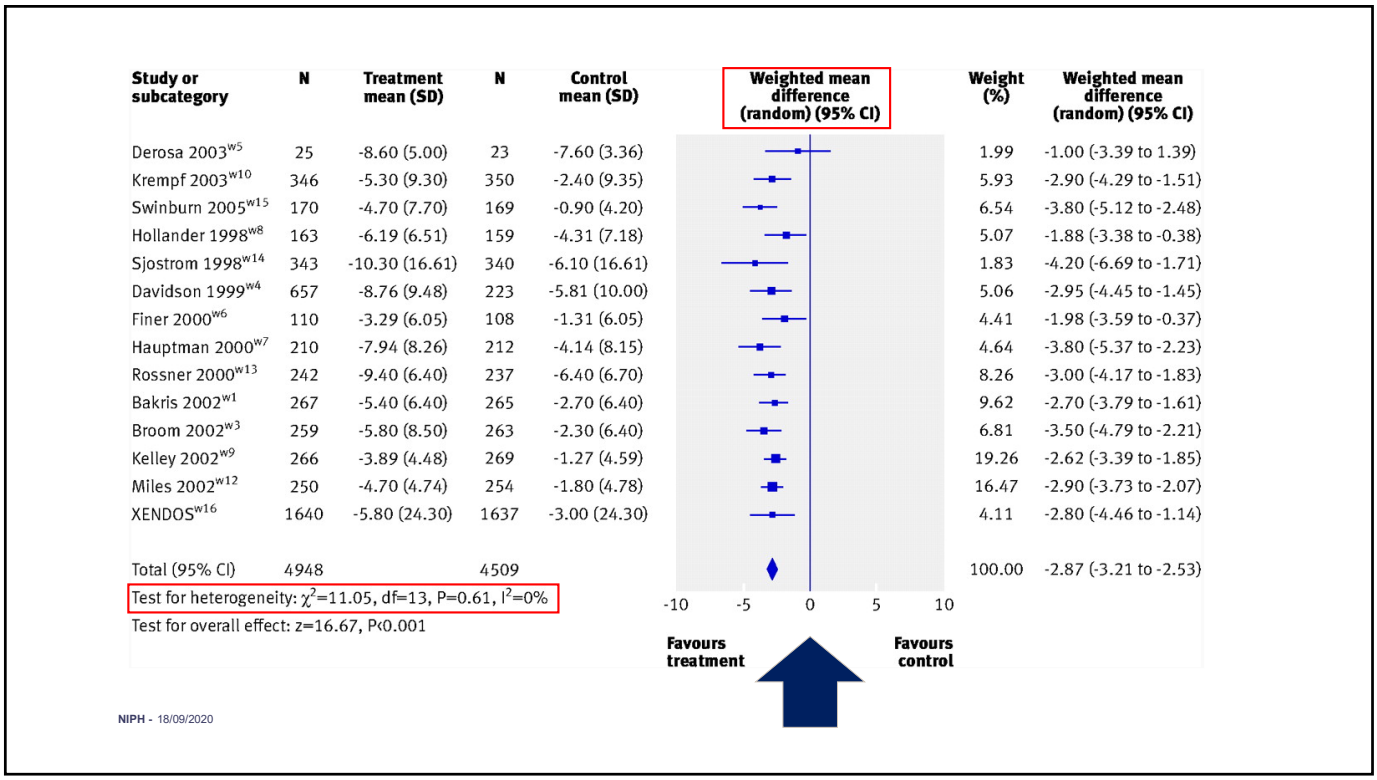
- The result of a meta-analysis is often showed in a forest plot
- A forest plot shows
  - The number of studies and participants
  - Which statistical model and effects measures that were chosen
  - The magnitude and direction of the effect in each study
  - The magnitude and direction of the effect in all studies when statistically combined
    - The effect estimate or population estimate
  - Measure/s of statistical heterogeneity
  - The weight of individual studies



26


No with recurrent headache/No in group
Study
Dexamethasone group
Placebo group
Relative risk (fixed) (95% CI)
Weight (%)
Relative risk (fixed) (95% CI)
Innes 1999[w1] 9/49 22/49 14.97 0.41 (0.21 to 0.80)
Jones 2003[w5] 8/34 10/36 6.61 0.85 (0.38 to 1.89)
Baden 2006[w2] 4/31 8/24 6.14 0.39 (0.13 to 1.13)
Donaldson 2006[w6] 21/57 18/42 14.10 0.86 (0.53 to 1.40)
Fiesseler 2006[w4] 19/44 20/41 14.09 0.89 (0.56 to 1.40)
Friedman 2007[w3] 39/106 43/99 30.26 0.85 (0.61 to 1.19)
Rowe 2007[w7] 14/64 20/62 13.83 0.68 (0.38 to 1.22)
Total (95% CI) 385 353 100.00 0.74 (0.60 to 0.90)
Test for heterogeneity: $\chi^2$=6.21, df=6, P=0.40, $I^2$=3.4%
Test for overall effect: z=3.01, P=0.003
0.1 0.2 0.5 1 2 5 10
Favours dexamethasone
Favours placebo
NIPH - 18/09/2020

27


Study or subcategory
N
Treatment mean (SD)
N
Control mean (SD)
Weighted mean difference (random) (95% CI)
Weight (%)
Weighted mean difference (random) (95% CI)
Derosa 2003[w5] 25 -8.60 (5.00) 23 -7.60 (3.36) 1.99 -1.00 (-3.39 to 1.39)
Krempf 2003[w10] 346 -5.30 (9.30) 350 -2.40 (9.35) 5.93 -2.90 (-4.29 to -1.51)
Swinburn 2005[w15] 170 -4.70 (7.70) 169 -0.90 (4.20) 6.54 -3.80 (-5.12 to -2.48)
Hollander 1998[w8] 163 -6.19 (6.51) 159 -4.31 (7.18) 5.07 -1.88 (-3.38 to -0.38)
Sjostrom 1998[w14] 343 -10.30 (16.61) 340 -6.10 (16.61) 1.83 -4.20 (-6.69 to -1.71)
Davidson 1999[w4] 657 -8.76 (9.48) 223 -5.81 (10.00) 5.06 -2.95 (-4.45 to -1.45)
Finer 2000[w6] 110 -3.29 (6.05) 108 -1.31 (6.05) 4.41 -1.98 (-3.59 to -0.37)
Hauptman 2000[w7] 210 -7.94 (8.26) 212 -4.14 (8.15) 4.64 -3.80 (-5.37 to -2.23)
Rossner 2000[w13] 242 -9.40 (6.40) 237 -6.40 (6.70) 8.26 -3.00 (-4.17 to -1.83)
Bakris 2002[w1] 267 -5.40 (6.40) 265 -2.70 (6.40) 9.62 -2.70 (-3.79 to -1.61)
Broom 2002[w3] 259 -5.80 (8.50) 263 -2.30 (6.40) 6.81 -3.50 (-4.79 to -2.21)
Kelley 2002[w9] 266 -3.89 (4.48) 269 -1.27 (4.59) 19.26 -2.62 (-3.39 to -1.85)
Miles 2002[w12] 250 -4.70 (4.74) 254 -1.80 (4.78) 16.47 -2.90 (-3.73 to -2.07)
XENDOS[w16] 1640 -5.80 (24.30) 1637 -3.00 (24.30) 4.11 -2.80 (-4.46 to -1.14)
Total (95% CI) 4948 4509 100.00 -2.87 (-3.21 to -2.53)
Test for heterogeneity: $\chi^2$=11.05, df=13, P=0.61, $I^2$=0%
Test for overall effect: z=16.67, P<0.001
-10 -5 0 5 10
Favours treatment
Favours control
NIPH - 18/09/2020

28

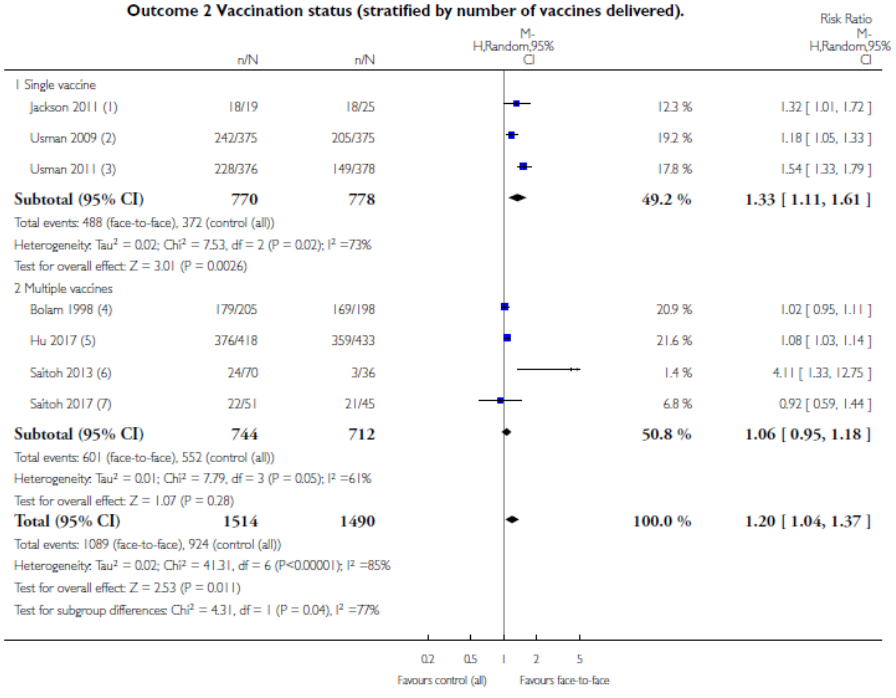
# Subgroup analysis

- Splitting all the participant data into subgroups
- May be done for subsets of participants (such as males and females), or for subsets of studies (such as different geographical locations)
- Investigate heterogeneous results
- Subgroup analyses are observational by nature and are not based on randomized comparisons!
- Findings may be misleading



29

**Analysis 1.2. Comparison 1 Face-to-face education versus control or non-face-to-face education (all), Outcome 2 Vaccination status (stratified by number of vaccines delivered).**






30

# Sensitivity analysis

- Are the results robust to the decisions made in the process of obtaining them?
  - Choices made in the process, e.g.
    - PICO
      - Characteristics of participants, interventions, comparators, outcomes
    - Data to be analysed
      - Missing data imputation, change scores/final values
    - Risk of bias
      - Low/high risk of bias



31

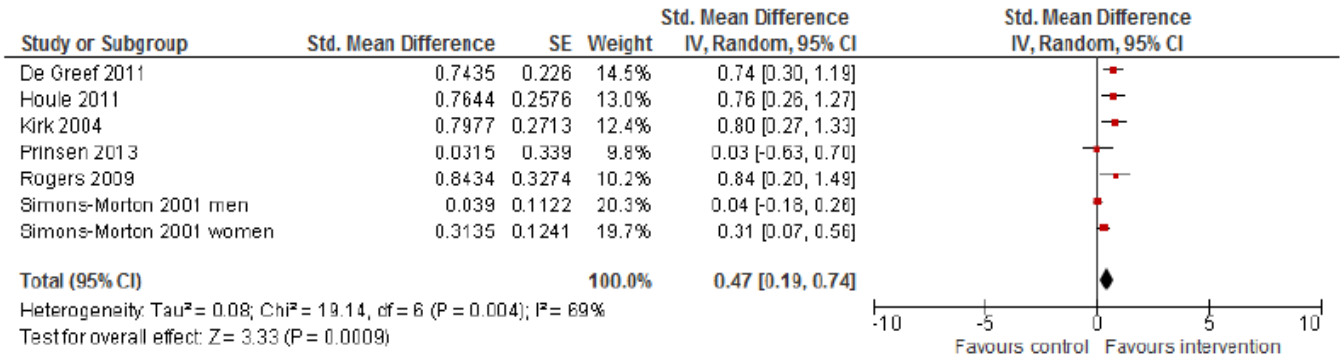



*Figure 4. Effects of cognitive therapies compared to no intervention or usual care on physical activity in studies reporting mean change difference.* SE = standard error, CI = confidence interval.



32

The pooled effect estimate was:
0.47 (0.19, 0.74)

Change in magnitude?

Change in direction of the effect?

**Table F1. Results of the sensitivity analyses in studies reporting mean change difference.**

| Study characteristic | SMD (95% CI) |
|---|---|
| **Health status** | |
| Patient group | 0.68 (0.43, 0.93) |
| Healthy with risk factor(s) for CVD | 0.17 (-0.10, 0.44) |
| **Duration of the intervention** | |
| 12 weeks | 0.57 (0.12, 1.03) |
| 36+ weeks | 0.41 (0.07, 0.74) |
| **Profession of person(s) delivering intervention** | |
| Psychologist + exercise specialist | 0.78 (0.41, 1.14) |
| Nurse/research assistant/health educator | 0.35 (0.06, 0.65) |
| **Risk of bias** | |
| Unclear | 0.64 (0.34, 0.94) |
| Low | 0.29 (-0.04, 0.62) |

SMD = standardized mean difference; CVD = cardiovascular disease



33

# Exercise 3 in break-out groups

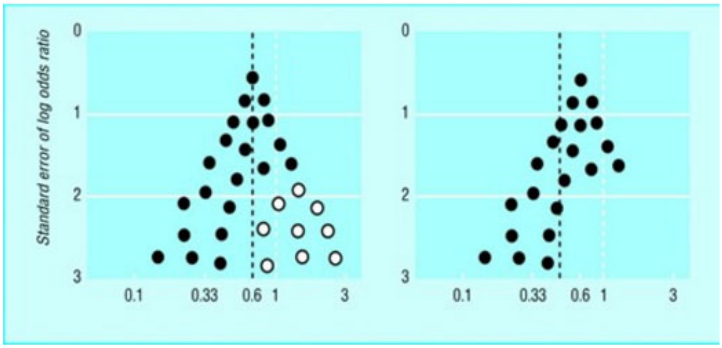
## Analyses and forest plots



34

# Publication bias

- Funnel plot






35

## Publication bias

| Table 5.8: Possible sources of publication bias throughout the publication process | |
|---|---|
| Phases of research publication | Actions contributing to or resulting in bias. |
| Preliminary and pilot studies | Small studies more likely to be "negative" (e.g. those with discarded or failed hypotheses) remain unpublished; companies classify some as proprietary information. |
| Report completion | Authors decide that reporting a "negative" study is uninteresting; and do not invest the time and effort required for submission. |
| Journal selection | Authors decide to submit the "negative" report to a nonindexed, non-English, or limited-circulation journal. |
| Editorial consideration | Editor decides that the "negative" study does not warrant peer review and rejects manuscript. |
| Peer review | Peer reviewers conclude that the "negative" study does not contribute to the field and recommend rejecting the manuscript. Author gives up or moves to lower impact journal. Publication delayed. |
| Author revision and resubmission | Author of rejected manuscript decides to forgo the submission of the "negative" study or to submit it again at a later time to another journal (see "journal selection" above). |
| Report publication | Journal delays the publication of the "negative" study.<br>Proprietary interests lead to report getting submitted to, and accepted by, different journals. |



36