MONASH University

Medicine, Nursing and Health Sciences

# *Challenges with continuous outcomes (part A)*

Jo McKenzie (joanne.mckenzie@monash.edu)
Cochrane Methods Training Event 2016
Birmingham, UK, 17-18th March 2016

# Session plan

1. Combining estimates from analyses of final values, change scores, and ANCOVA

   *Practical (group discussion)*

2. Meta-analysis of skewed data

   *Practical (computer)*

3. Ratio of arithmetic means

   *Practical (computer)*

# Effect measures for continuous outcomes

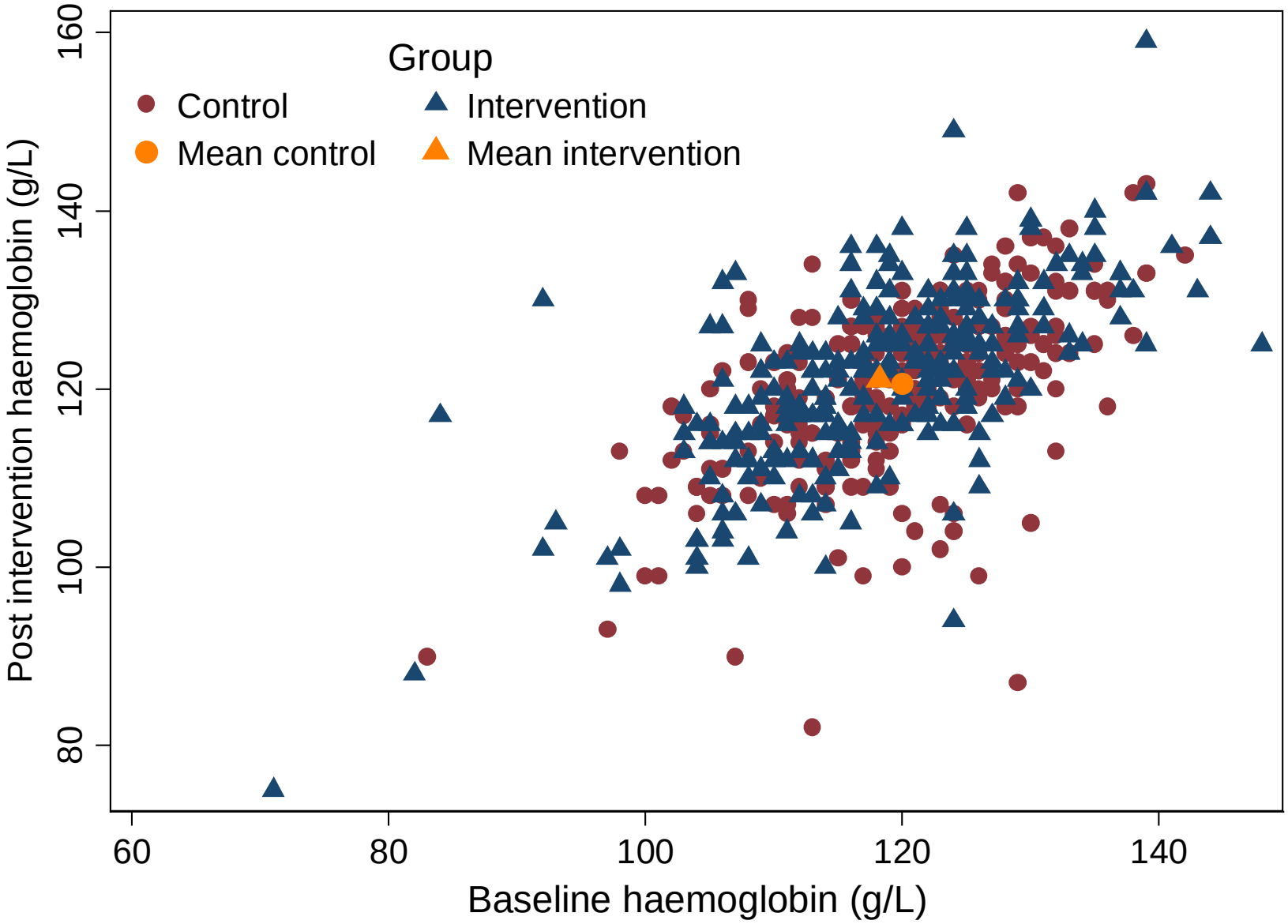
| Effect measure | |
|---|---|
| *Difference effect measures* | |
| Mean difference | $MD = \bar{Y}_{int} - \bar{Y}_{ctrl}$ |
| Standardised mean difference | $SMD = \left( \frac{\bar{Y}_{int} - \bar{Y}_{ctrl}}{S_{pooled}} \right)$ |
| *Ratio effect measures* | |
| Ratio of means | $RoM = \frac{\bar{Y}_{int}}{\bar{Y}_{ctrl}}$ |
| Ratio of geometric means | $RoGM = \frac{G_{int}}{G_{ctrl}}$ |

# Combining estimates from analyses of final values, change scores, and ANCOVA

# Analysis of a randomised trial with measurements before and after intervention

- Randomised trial carried out in the Ubon Ratchathani province NE Thailand
- Aimed to test the efficacy of a seasoning powder fortified with micronutrients
- Groups:
  - Intervention: fortified seasoning powder added to instant wheat noodles or rice
  - Control: unfortified seasoning powder added to instant wheat noodles or rice
- Data collected at baseline and follow-up (31 weeks)
- Primary outcome was anaemia (defined from the continuous variable haemoglobin)

Post intervention haemoglobin vs baseline haemoglobin



# Analysis options

- Ignore the baseline values, and calculate the difference in means at follow-up between groups (simple analysis of final values SAFV)

$$\hat{\theta}_{SAFV} = \bar{Y}_{int} - \bar{Y}_{ctrl}$$

- Adjust for baseline by calculating the difference in mean change $(Y - X)$ between groups (simple analysis of change scores SACS)

$$\hat{\theta}_{SACS} = (\bar{Y}_{int} - \bar{Y}_{ctrl}) - (\bar{X}_{int} - \bar{X}_{ctrl})$$

- Adjust for baseline using regression modelling (ANCOVA)

$$\hat{\theta}_{ANCOVA} = (\bar{Y}_{int} - \bar{Y}_{ctrl}) - \hat{\beta}(\bar{X}_{int} - \bar{X}_{ctrl})$$

where $\beta = \rho \frac{\sigma_Y}{\sigma_X}$ and is estimated from the regression model

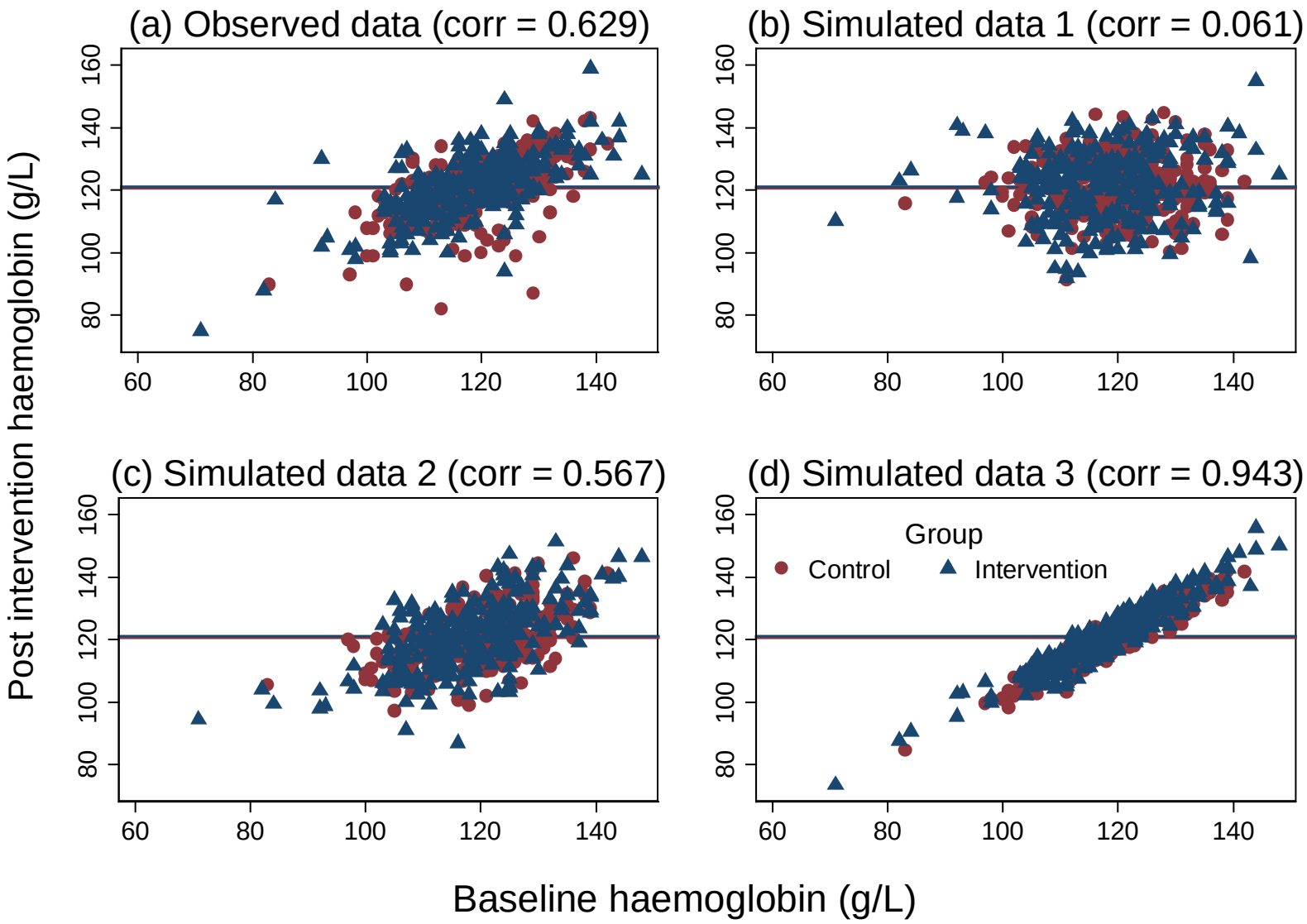
# Data sets used to illustrate the impact of correlation on intervention effect estimates

| Dataset | Observed correlation | Follow-up haemoglobin (g/L) | | | |
|---|---|---|---|---|---|
| | | Intervention group | | Control group | |
| | | Mean | SD | Mean | SD |
| Observed data | 0.629 | 121.0 | 10.1 | 120.5 | 9.5 |

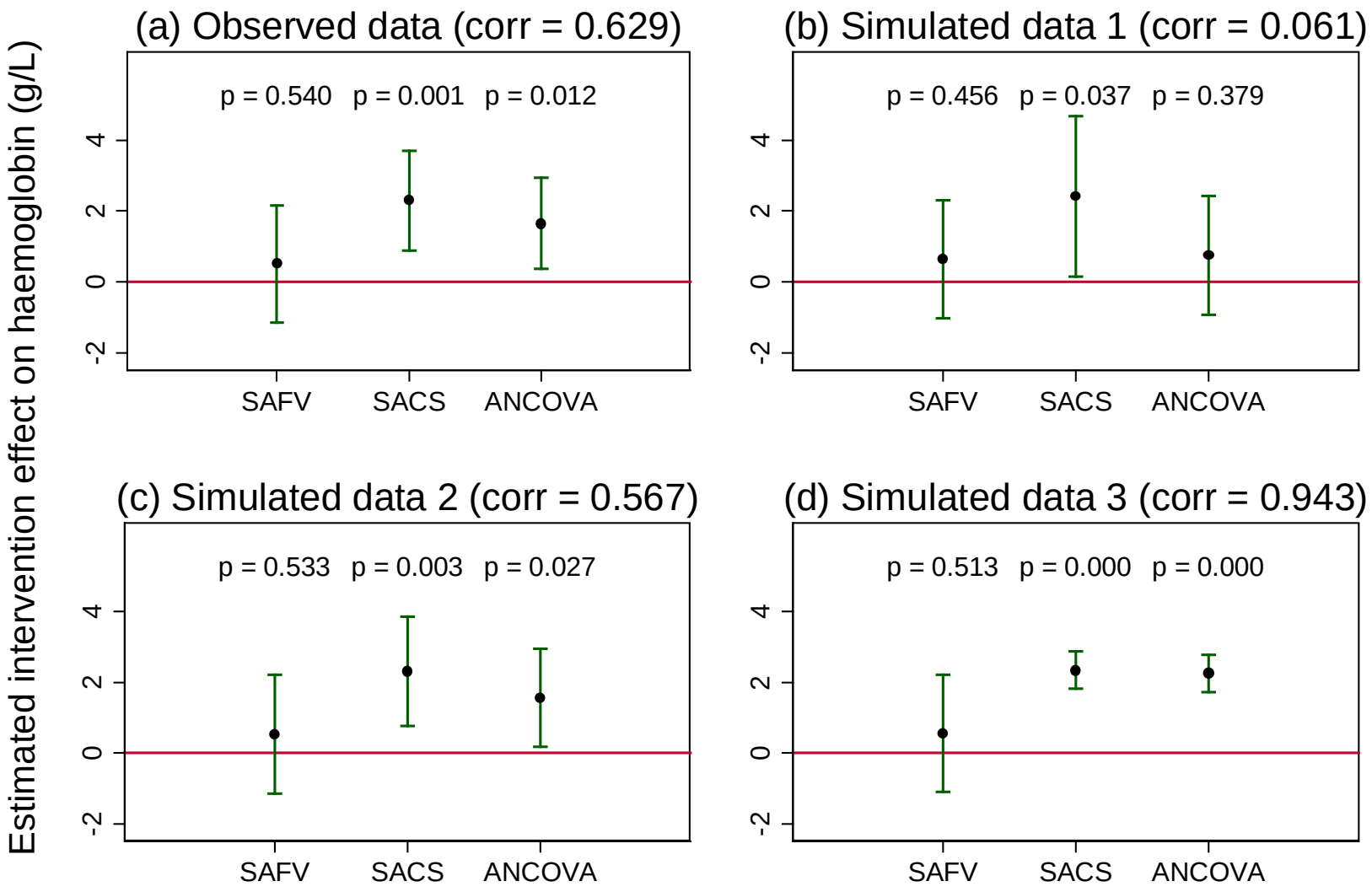
# Data sets used to illustrate the impact of correlation on intervention effect estimates

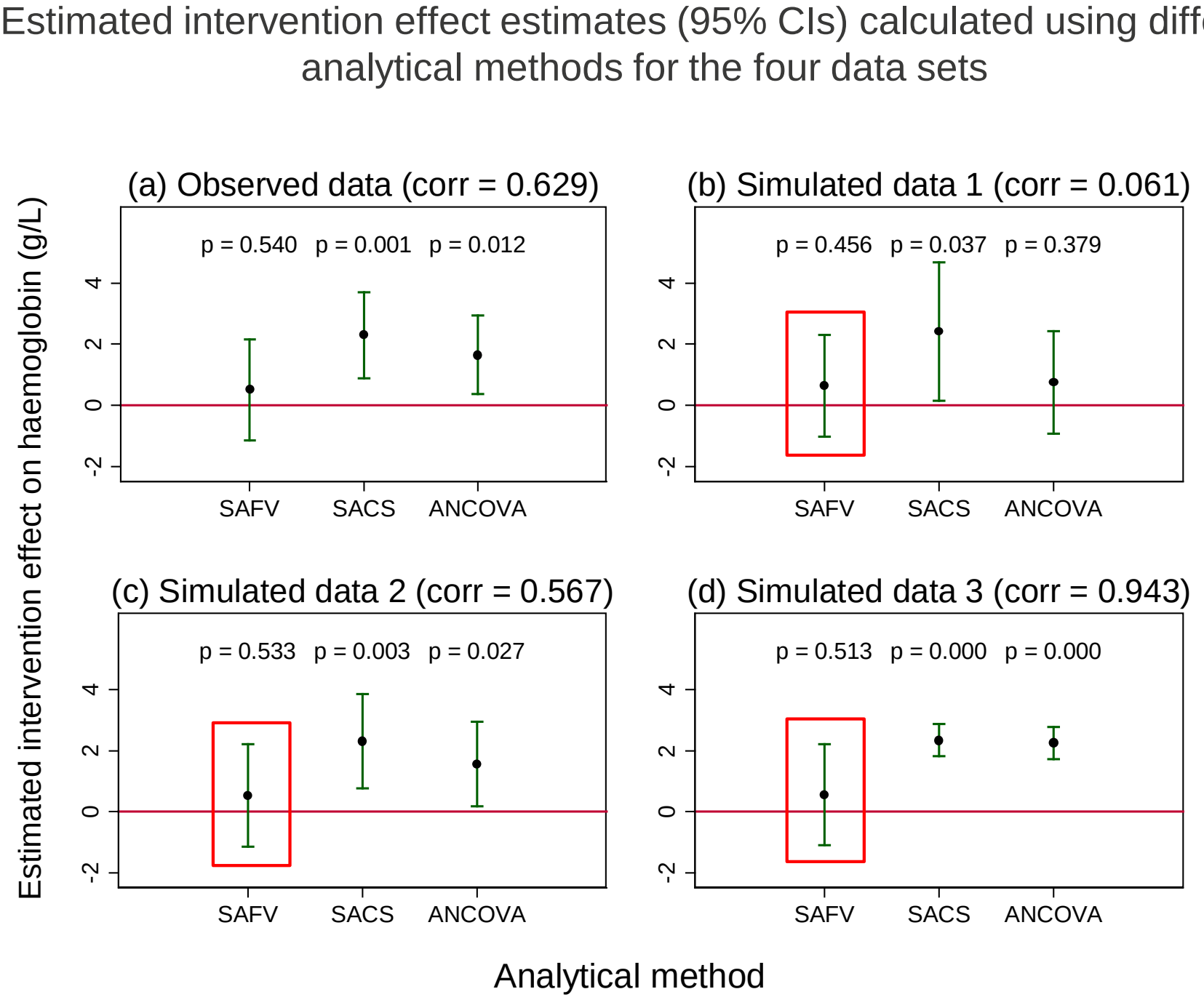
| Dataset | Observed correlation | Follow-up haemoglobin (g/L) | | | |
|---|---|---|---|---|---|
| | | Intervention group | | Control group | |
| | | Mean | SD | Mean | SD |
| Observed data | 0.629 | 121.0 | 10.1 | 120.5 | 9.5 |
| Simulated data 1 | 0.061 | 121.2 | 10.8 | 120.6 | 8.8 |
| Simulated data 2 | 0.567 | 121.2 | 10.8 | 120.6 | 8.8 |
| Simulated data 3 | 0.943 | 121.1 | 10.5 | 120.5 | 9.0 |

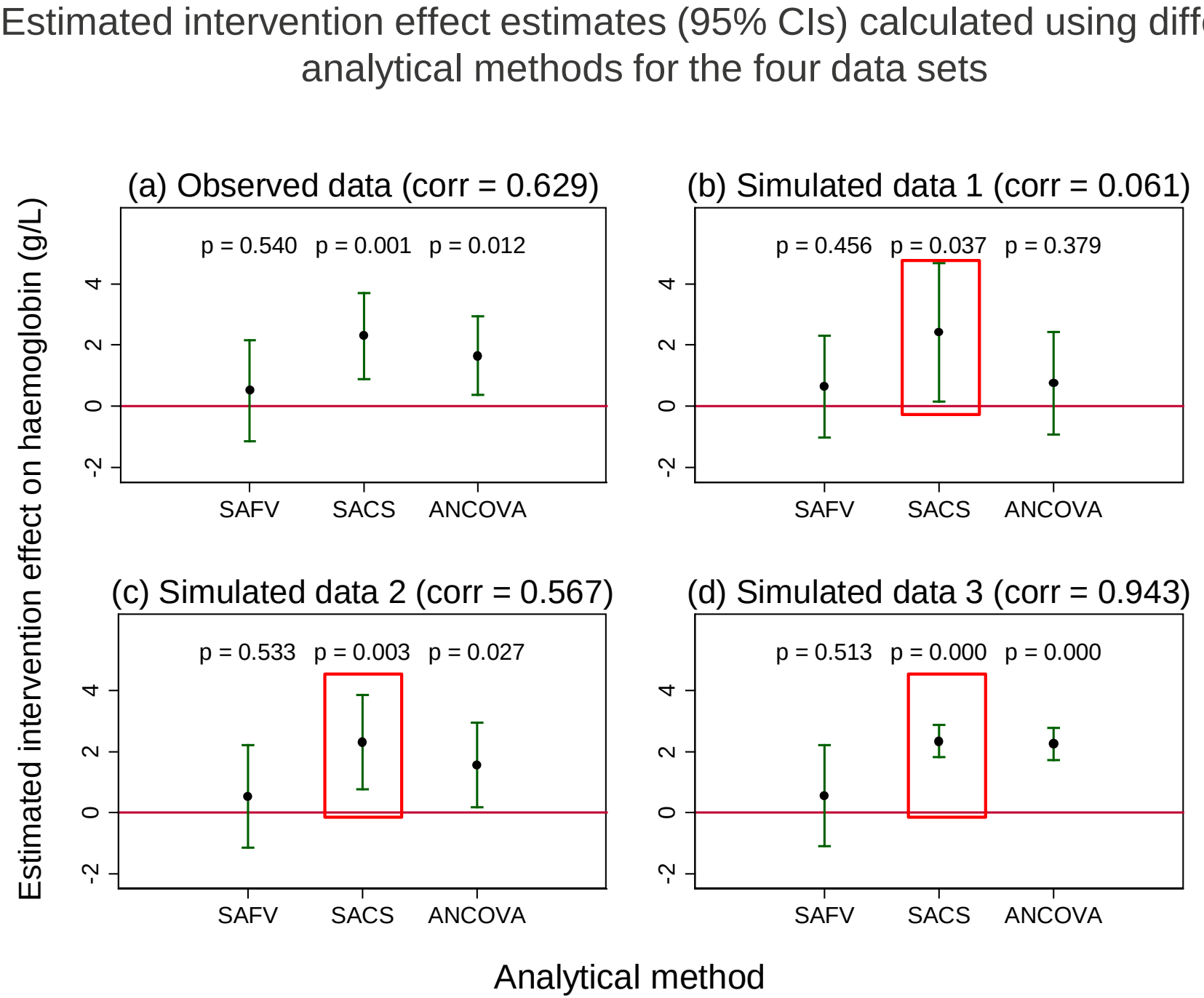
Scatter plots of post intervention haemoglobin vs baseline haemoglobin for observed and simulated data sets

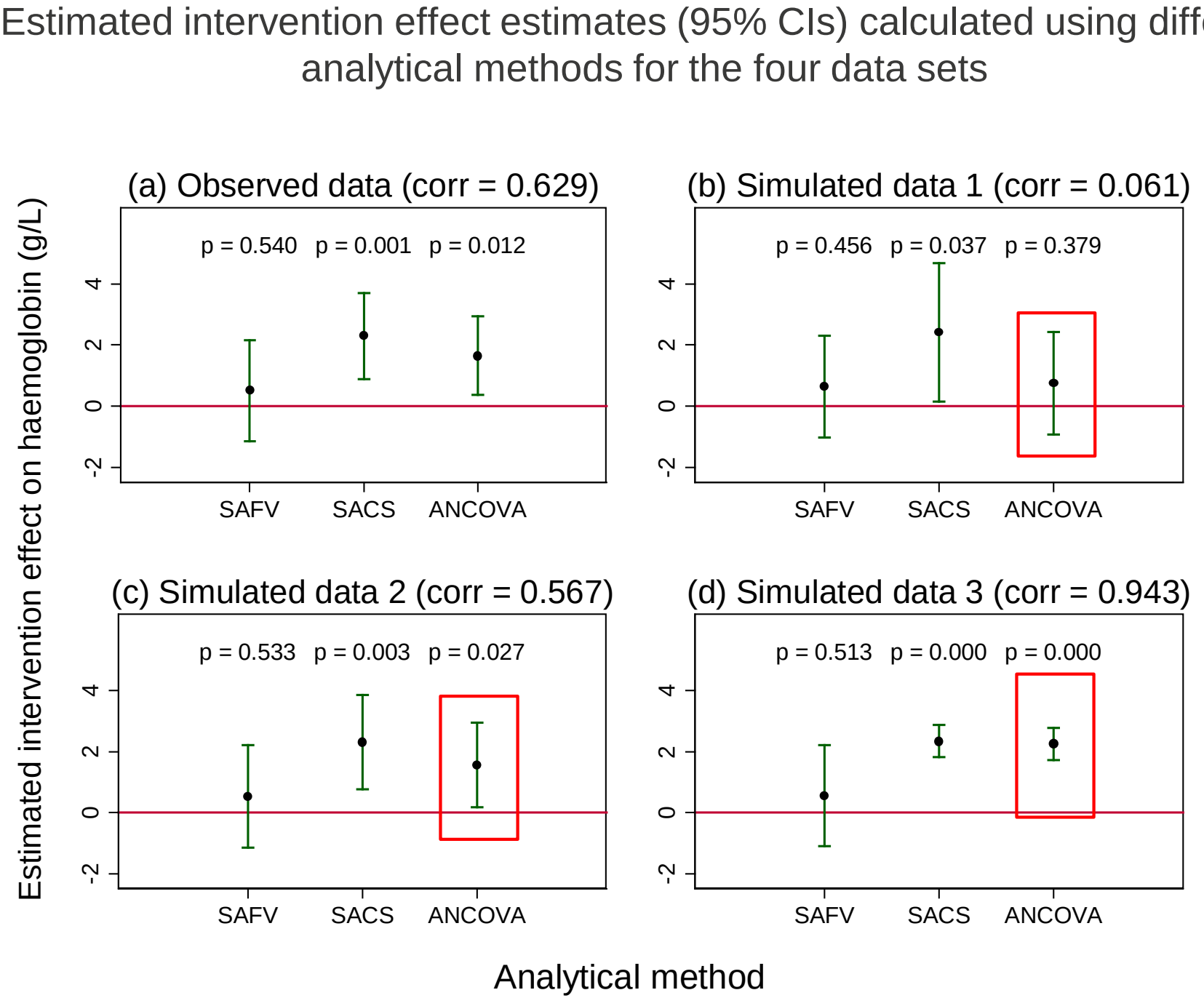


# Estimated intervention effect estimates (95% CIs) calculated using different analytical methods for the four data sets




Estimated intervention effect estimates (95% CIs) calculated using different analytical methods for the four data sets
(a) Observed data (corr = 0.629)
p = 0.540 p = 0.001 p = 0.012
(b) Simulated data 1 (corr = 0.061)
p = 0.456 p = 0.037 p = 0.379
(c) Simulated data 2 (corr = 0.567)
p = 0.533 p = 0.003 p = 0.027
(d) Simulated data 3 (corr = 0.943)
p = 0.513 p = 0.000 p = 0.000
SAFV
SACS
ANCOVA
-2
0
2
4
Estimated intervention effect on haemoglobin (g/L)
Analytical method


Estimated intervention effect estimates (95% CIs) calculated using different analytical methods for the four data sets
(a) Observed data (corr = 0.629)
p = 0.540 p = 0.001 p = 0.012
(b) Simulated data 1 (corr = 0.061)
p = 0.456 p = 0.037 p = 0.379
(c) Simulated data 2 (corr = 0.567)
p = 0.533 p = 0.003 p = 0.027
(d) Simulated data 3 (corr = 0.943)
p = 0.513 p = 0.000 p = 0.000
SAFV
SACS
ANCOVA
-2
0
2
4
Estimated intervention effect on haemoglobin (g/L)
Analytical method


Estimated intervention effect estimates (95% CIs) calculated using different analytical methods for the four data sets
(a) Observed data (corr = 0.629)
p = 0.540 p = 0.001 p = 0.012
(b) Simulated data 1 (corr = 0.061)
p = 0.456 p = 0.037 p = 0.379
(c) Simulated data 2 (corr = 0.567)
p = 0.533 p = 0.003 p = 0.027
(d) Simulated data 3 (corr = 0.943)
p = 0.513 p = 0.000 p = 0.000
SAFV
SACS
ANCOVA
-2
0
2
4
Estimated intervention effect on haemoglobin (g/L)
Analytical method

# Comparing the trial analysis methods

- Estimates of intervention effect:
  - For a particular data set, the three analytical methods can produce different estimates of intervention effect
  - Over the data sets (varying correlation), the ANCOVA estimate varies; SACS or SAFV estimates do not
- Standard errors:
  - The SE of the SAFV estimate is not affected by correlation
  - Increasing correlation results in a smaller SE for the SACS estimate
  - Correlation < 0.5, the SE of SACS estimate is > SE of the FV estimate. This is reversed when the correlation is > 0.5
  - For a particular correlation, the SE of the ANCOVA estimate is smaller compared with SEs of FV and CS estimates

# Relationship between the three analysis methods

$$\hat{\theta}_{ANCOVA} = (\bar{Y}_{int} - \bar{Y}_{ctrl}) - \rho(\bar{X}_{int} - \bar{X}_{ctrl})$$

(assuming $\sigma_Y^2 = \sigma_X^2$)

- Scenario 1: $\rho$ is close to 0

$$\hat{\theta}_{ANCOVA} \approx \hat{\theta}_{SAFV}$$

- Scenario 2: $\rho$ is close to 1

$$\hat{\theta}_{ANCOVA} \approx \hat{\theta}_{SACS}$$

- Scenario 3: minimal baseline imbalance, i.e. $(\bar{X}_{int} - \bar{X}_{ctrl}) \approx 0$

$$\hat{\theta}_{ANCOVA} \approx \hat{\theta}_{SACS} \approx \hat{\theta}_{SAFV}$$

# *Practical (group discussion)*

# When undertaking a systematic review …

- Likely to encounter estimates calculated from different analysis methods
- This could include:
  - SACS
  - SAFV
  - Sometimes SACS *and* SAFV
  - Sometimes ANCOVA

- *Practical:* Discussion of a meta-analysis from a systematic review examining the effect of calcium supplementation on body weight (Trowman 2006 Br J Nut)

# Study characteristics (modified table 1) (Trowman 2006 Br J Nut)

| Study | Number of participants | Age* | Sex | Intervention (Ca concentration) | Length of follow-up | Country |
|---|---|---|---|---|---|---|
| Chee et al. (2003) | 173 | 58·9 | Female (postmenopausal) | Ca supplement (1200 mg/d) | 24 months | Malaysia |
| Jensen et al. (2001) | 52 | NA | Female (obese postmenopausal) | Ca supplement (1000 mg/d) | 26 weeks | Denmark |
| Lau et al. (2001) | 185 | 57·0 | Female (postmenopausal) | Ca supplement (800 mg/d) | 24 months | China |
| Reid et al. (2002) | 223 | 72·0 | Female (postmenopausal) | Ca supplement (1000 mg/d) | 24 months | New Zealand |
| Shapses et al. (2004) | 36 | 59·3 | Female (obese postmenopausal) | Ca supplement (1000 mg/d) | 25 weeks | USA |
| Shapses et al. (2004) | 30 | 56·0 | Female (obese postmenopausal) | Ca supplement (1000 mg/d) | 25 weeks | USA |
| Shapses et al. (2004) | 42 | 41·0 | Female (obese postmenopausal) | Ca supplement (1000 mg/d) | 25 weeks | USA |
| Winters-Stone & Snow (2004) | 23 | 24·8 | Female (athletes) | Ca supplement (1000 mg/d) | 12 months | USA |
| Zemel et al. (2004) | 41 | 46 | Mixed (obese) | Calcium supplement (800 mg/d) | 24 weeks | USA |

**NA, not available**
*** Mean age. When age was reported separately by subgroups, the mean between the groups was calculated.**

# Calcium supplementation on body weight (Trowman 2006 Br J Nut)

| Trial | Year | Baseline (weight kg) | | | | Follow-up (weight kg) | | Change (weight kg) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Intervention | | Control | | Intervention | Control | Intervention | Control |
| | | N | Mean (SD) | n | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) |
| Chee | 2003 | 91 | 56.1 (8.9) | 82 | 57.2 (9.4) | | | 0.0 (2.6) [a] | 0.2 (2.6) [a] |
| Jensen | 2001 | 25 | 94.6 (14.0)[a] | 27 | 93.8 (14.0)[a] | 89.0 (12.7)[a] | 89.1 (14.7)[a] | | |
| Lau | 2001 | 95 | 56.9 (7.1) | 90 | 58.9 (7.5) | | | 0.5 (2.6)[a] | -0.3 (2.7)[a] |
| Reid | 2002 | 111 | 66.0 (10.0) | 112 | 68.0 (11.0) | | | -0.3 (1.8) | -0.1 (2.4) |
| Shapses 1[c] | 2004 | 17 | 84.1 (9.4) | 19 | 89.4 (10.3) | | | -7.0 (4.6) | -7.3 (5.3) |
| Shapses 2[c] | 2004 | 11 | 85.9 (9.2) | 11 | 94.2 (15.7) | | | -6.7 (2.6) | -7.6 (5.7) |
| Shapses 3[c] | 2004 | 18 | 93.7 (13.6) | 24 | 93.5 (14.3) | | | -6.7 (5.5) | -4.3 (3.5) |
| Winters-Stone | 2004 | 13 | 57.2 (4.9) | 10 | 54.1 (7.2) | 56.3 (4.3) | 54.8 (7.2) | | |
| Zemel | 2004 | 11 | 99.8 (14.9) | 10 | 103.1 (19.3) | | | -8.6 (5.3)[a] | -6.6 (8.2) [a] |

a Calculated from the standard error
b Follow-up sample size ntrt = 24 and nctrl = 24
c Shapses *et al* (Shapses *et al*, 2004) report on three randomised controlled trials.
Trials 1, 2, and 3 include postmenopausal women, postmenopausal women special diet, and premenopausal women respectively.

# Practical

- How would you undertake a meta-analysis of this data?
  - What data would you choose, and why?
  - Would you impute any data, and how?

# Meta-analysis options: a proposed hierarchy

| Option | Advantages | Disadvantages |
|---|---|---|
| **1) Individual patient data:** Obtain IPD for each trial. Reanalyse using conventional two-step approach, or a more complex approach such as multilevel modelling. ANCOVA would be the method of choice within each trial | • Avoids the issue of trialists selectively reporting results<br>• Able to re-analyse data in a consistent way<br>• Potentially adjust for other prognostic factors<br>• Can use the most powerful analytical method<br>• Do not have to rely on summary data provided in publications | • Generally not possible to obtain IPD |

# Meta-analysis options: a proposed hierarchy

| Option | Advantages | Disadvantages |
|---|---|---|
| **1) Individual patient data:** Obtain IPD for each trial. Reanalyse using conventional two-step approach, or a more complex approach such as multilevel modelling. ANCOVA would be the method of choice within each trial | • Avoids the issue of trialists selectively reporting results<br>• Able to re-analyse data in a consistent way<br>• Potentially adjust for other prognostic factors<br>• Can use the most powerful analytical method<br>• Do not have to rely on summary data provided in publications | • Generally not possible to obtain IPD |
| **2) Meta-analysis using only ANCOVA results:** Use available ANCOVA estimates. When not available, recreate the estimates from available summary statistics, or imputing missing statistics (e.g. correlations) | • Reduce bias from random baseline imbalance across the included randomised trials or from selective reporting of results<br>• May provide greater precision compared with pooling results from SAFV or SACS | • Will generally require assumptions to be made regarding the correlation<br>• Will require data manipulation |

# Meta-analysis options: a proposed hierarchy

| Option | Advantages | Disadvantages |
|---|---|---|
| **3) Meta-analysis using results from only one analysis method (SAFV or SACS):** Meta-analyse estimates using the same analysis method. This approach may involve imputing missing statistics (e.g. correlations) | • Removes bias from trialists selectively reporting analyses | • Can provide a biased pooled estimate when there is baseline imbalance across randomised trials. Generally only a problem when there are a small number of trials with few participants<br>• May provide less precision compared with meta-analysing ANCOVA results<br>• May require assumptions to be made about missing data<br>• Will require data manipulation (generally less than option 2) |

# Meta-analysis options: a proposed hierarchy

| Option | Advantages | Disadvantages |
|---|---|---|
| **3) Meta-analysis using results from only one analysis method (SAFV or SACS):** Meta-analyse estimates using the same analysis method. This approach may involve imputing missing statistics (e.g. correlations) | • Removes bias from trialists selectively reporting analyses | • Can provide a biased pooled estimate when there is baseline imbalance across randomised trials. Generally only a problem when there are a small number of trials with few participants<br>• May provide less precision compared with meta-analysing ANCOVA results<br>• May require assumptions to be made about missing data<br>• Will require data manipulation (generally less than option 2) |
| **4) Meta-analysis using a mix of results from different analysis methods:** The meta-analysis may include estimates from SAFV, SACS, and ANCOVA | • Generally less imputation and data manipulation required | • Prone to selective reporting of results<br>• May provide less precision compared with meta-analysing ANCOVA results |

# Calcium supplementation on body weight (Trowman 2006 Br J Nut)

| Trial | Year | Baseline (weight kg) | | | | Follow-up (weight kg) | | Change (weight kg) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Intervention | | Control | | Intervention | Control | Intervention | Control |
| | | n | Mean (SD) | n | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) |
| Chee | 2003 | 91 | 56.1 (8.9) | 82 | 57.2 (9.4) | *56.1* ***(?)*** | *57.4* ***(?)*** | 0.0 (2.6) [a] | 0.2 (2.6) [a] |
| Jensen | 2001 | 25 | 94.6 (14.0)[a] | 27 | 93.8 (14.0)[a] | 89.0 (12.7)[a] | 89.1 (14.7)[a] | *-5.6* ***(?)*** | *-4.7* ***(?)*** |
| Lau | 2001 | 95 | 56.9 (7.1) | 90 | 58.9 (7.5) | *57.4* ***(?)*** | *58.6* ***(?)*** | 0.5 (2.6)[a] | -0.3 (2.7)[a] |
| Reid | 2002 | 111 | 66.0 (10.0) | 112 | 68.0 (11.0) | *65.7* ***(?)*** | *67.9* ***(?)*** | -0.3 (1.8) | -0.1 (2.4) |
| Shapses 1[c] | 2004 | 17 | 84.1 (9.4) | 19 | 89.4 (10.3) | *77.1* ***(?)*** | *82.1* ***(?)*** | -7.0 (4.6) | -7.3 (5.3) |
| Shapses 2[c] | 2004 | 11 | 85.9 (9.2) | 11 | 94.2 (15.7) | *79.2* ***(?)*** | *86.6* ***(?)*** | -6.7 (2.6) | -7.6 (5.7) |
| Shapses 3[c] | 2004 | 18 | 93.7 (13.6) | 24 | 93.5 (14.3) | *87.0* ***(?)*** | *89.2* ***(?)*** | -6.7 (5.5) | -4.3 (3.5) |
| Winters-Stone | 2004 | 13 | 57.2 (4.9) | 10 | 54.1 (7.2) | 56.3 (4.3) | 54.8 (7.2) | *-0.9* ***(?)*** | *0.7* ***(?)*** |
| Zemel | 2004 | 11 | 99.8 (14.9) | 10 | 103.1 (19.3) | *91.2* ***(?)*** | *96.5* ***(?)*** | -8.6 (5.3)[a] | -6.6 (8.2) [a] |

a Calculated from the standard error
b Follow-up sample size ntrt = 24 and nctrl = 24
c Shapses *et al* (Shapses *et al*, 2004) report on three randomised controlled trials.
Trials 1, 2, and 3 include postmenopausal women, postmenopausal women special diet, and premenopausal women respectively.

## Option 3 (v1): Meta-analysis of results from only one analysis method

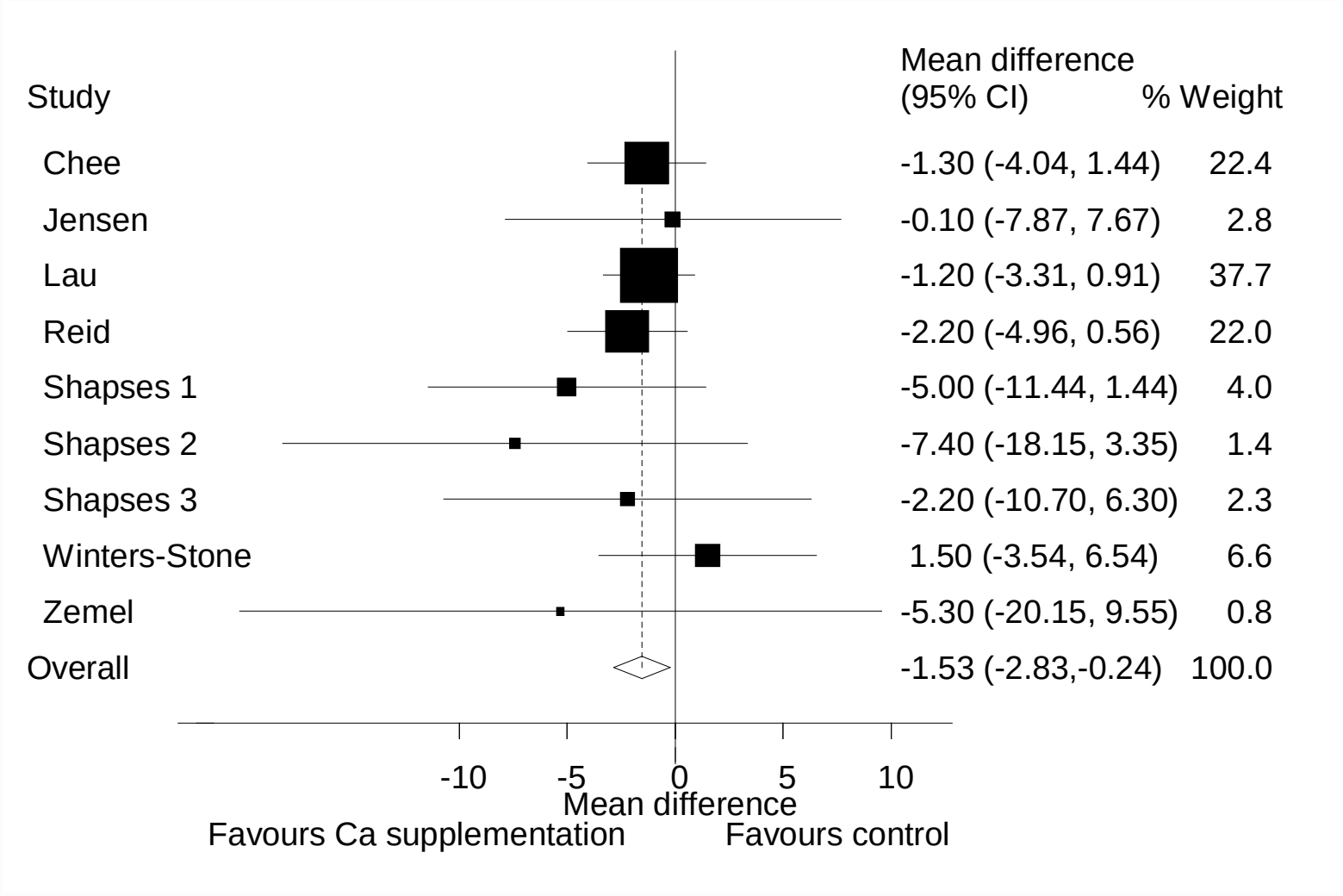
- Trowman (2006) used this option
- For each trial, estimated treatment effect by calculating the difference in mean follow-up measurements
- For missing SDs at follow-up, they assumed the baseline SD

# Calcium supplementation on body weight (Trowman 2006 Br J Nut)

| Trial | Year | Baseline (weight kg) | | | | Follow-up (weight kg) | | Change (weight kg) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Intervention | | Control | | Intervention | Control | Intervention | Control |
| | | n | Mean (SD) | n | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) |
| Chee | 2003 | 91 | 56.1 (8.9) | 82 | 57.2 (9.4) | *56.1* ***(8.9)*** | *57.4* ***(9.4)*** | 0.0 (2.6) [a] | 0.2 (2.6) [a] |
| Jensen | 2001 | 25 | 94.6 (14.0)[a] | 27 | 93.8 (14.0)[a] | 89.0 (12.7)[a] | 89.1 (14.7)[a] | *-5.6* ***(?)*** | *-4.7* ***(?)*** |
| Lau | 2001 | 95 | 56.9 (7.1) | 90 | 58.9 (7.5) | *57.4* ***(7.1)*** | *58.6* ***(7.5)*** | 0.5 (2.6)[a] | -0.3 (2.7)[a] |
| Reid | 2002 | 111 | 66.0 (10.0) | 112 | 68.0 (11.0) | *65.7* ***(10.0)*** | *67.9* ***(11.0)*** | -0.3 (1.8) | -0.1 (2.4) |
| Shapses 1[c] | 2004 | 17 | 84.1 (9.4) | 19 | 89.4 (10.3) | *77.1* ***(9.4)*** | *82.1* ***(10.3)*** | -7.0 (4.6) | -7.3 (5.3) |
| Shapses 2[c] | 2004 | 11 | 85.9 (9.2) | 11 | 94.2 (15.7) | *79.2* ***(9.2)*** | *86.6* ***(15.7)*** | -6.7 (2.6) | -7.6 (5.7) |
| Shapses 3[c] | 2004 | 18 | 93.7 (13.6) | 24 | 93.5 (14.3) | *87.0* ***(13.6)*** | *89.2* ***(14.3)*** | -6.7 (5.5) | -4.3 (3.5) |
| Winters-Stone | 2004 | 13 | 57.2 (4.9) | 10 | 54.1 (7.2) | 56.3 (4.3) | 54.8 (7.2) | *-0.9* ***(?)*** | *0.7* ***(?)*** |
| Zemel | 2004 | 11 | 99.8 (14.9) | 10 | 103.1 (19.3) | *91.2* ***(14.9)*** | *96.5* ***(19.3)*** | -8.6 (5.3)[a] | -6.6 (8.2) [a] |

a Calculated from the standard error
b Follow-up sample size ntrt = 24 and nctrl = 24
c Shapses *et al* (Shapses *et al*, 2004) report on three randomised controlled trials.
Trials 1, 2, and 3 include postmenopausal women, postmenopausal women special diet, and premenopausal women respectively.

# Combining intervention estimates from SAFV only



## Option 3 (v2): Meta-analysis of results from only one analysis method

- Use change scores
  - Impute missing change score SDs, in each intervention group, by calculating the median of the other SDs
- Many other options for imputing missing SDs

  e.g. in trials with baseline SDs and change SDs, assume follow-up SDs are the same as baseline SDs (seems reasonable assumption based on Jensen 2001 and Winters-Stone 2004) then calculate correlations and SDs at follow-up using

$$Corr_{int} = \frac{SD_{int,X}^2 + SD_{int,Y}^2 - SD_{int,C}^2}{2 \times SD_{int,X} \times SD_{int,Y}}$$

$$SD_{int,C} = \sqrt{SD_{int,X}^2 + SD_{int,Y}^2 - \left(2corr_{int} \times SD_{int,X} \times SD_{int,Y}\right)}$$

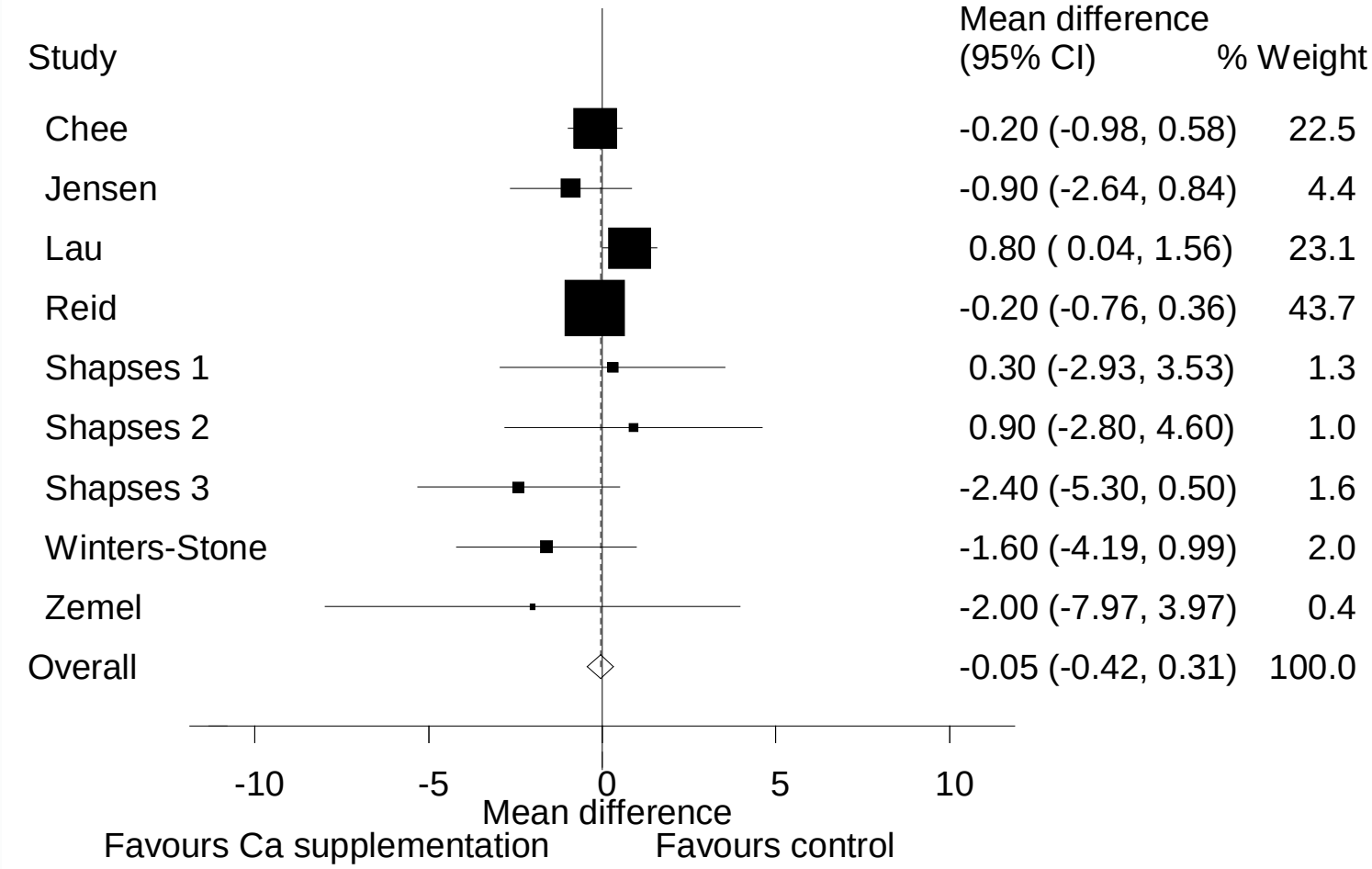
***Wiebe 2006 J Clin Epi; Balk 2012 AHRQ***

# Calcium supplementation on body weight (Trowman 2006 Br J Nut)

| Trial | Year | Baseline (weight kg) | | | | Follow-up (weight kg) | | Change (weight kg) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Intervention | | Control | | Intervention | Control | Intervention | Control |
| | | N | Mean (SD) | n | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) |
| Chee | 2003 | 91 | 56.1 (8.9) | 82 | 57.2 (9.4) | *56.1* | *57.4* | 0.0 (2.6) [a] | 0.2 (2.6) [a] |
| Jensen | 2001 | 25 | 94.6 (14.0)[a] | 27 | 93.8 (14.0)[a] | 89.0 (12.7)[a] | 89.1 (14.7)[a] | *-5.6* ***(2.6)*** | *-4.7* ***(3.5)*** |
| Lau | 2001 | 95 | 56.9 (7.1) | 90 | 58.9 (7.5) | *57.4* | *58.6* | 0.5 (2.6)[a] | -0.3 (2.7)[a] |
| Reid | 2002 | 111 | 66.0 (10.0) | 112 | 68.0 (11.0) | *65.7* | *67.9* | -0.3 (1.8) | -0.1 (2.4) |
| Shapses 1[c] | 2004 | 17 | 84.1 (9.4) | 19 | 89.4 (10.3) | *77.1* | *82.1* | -7.0 (4.6) | -7.3 (5.3) |
| Shapses 2[c] | 2004 | 11 | 85.9 (9.2) | 11 | 94.2 (15.7) | *79.2* | *86.6* | -6.7 (2.6) | -7.6 (5.7) |
| Shapses 3[c] | 2004 | 18 | 93.7 (13.6) | 24 | 93.5 (14.3) | *87.0* | *89.2* | -6.7 (5.5) | -4.3 (3.5) |
| Winters-Stone | 2004 | 13 | 57.2 (4.9) | 10 | 54.1 (7.2) | 56.3 (4.3) | 54.8 (7.2) | *-0.9* ***(2.6)*** | *0.7* ***(3.5)*** |
| Zemel | 2004 | 11 | 99.8 (14.9) | 10 | 103.1 (19.3) | *91.2* | *96.5* | -8.6 (5.3)[a] | -6.6 (8.2) [a] |

a Calculated from the standard error
b Follow-up sample size ntrt = 24 and nctrl = 24
c Shapses *et al* (Shapses *et al*, 2004) report on three randomised controlled trials.
Trials 1, 2, and 3 include postmenopausal women, postmenopausal women special diet, and premenopausal women respectively.

# Combining intervention estimates from SACS only



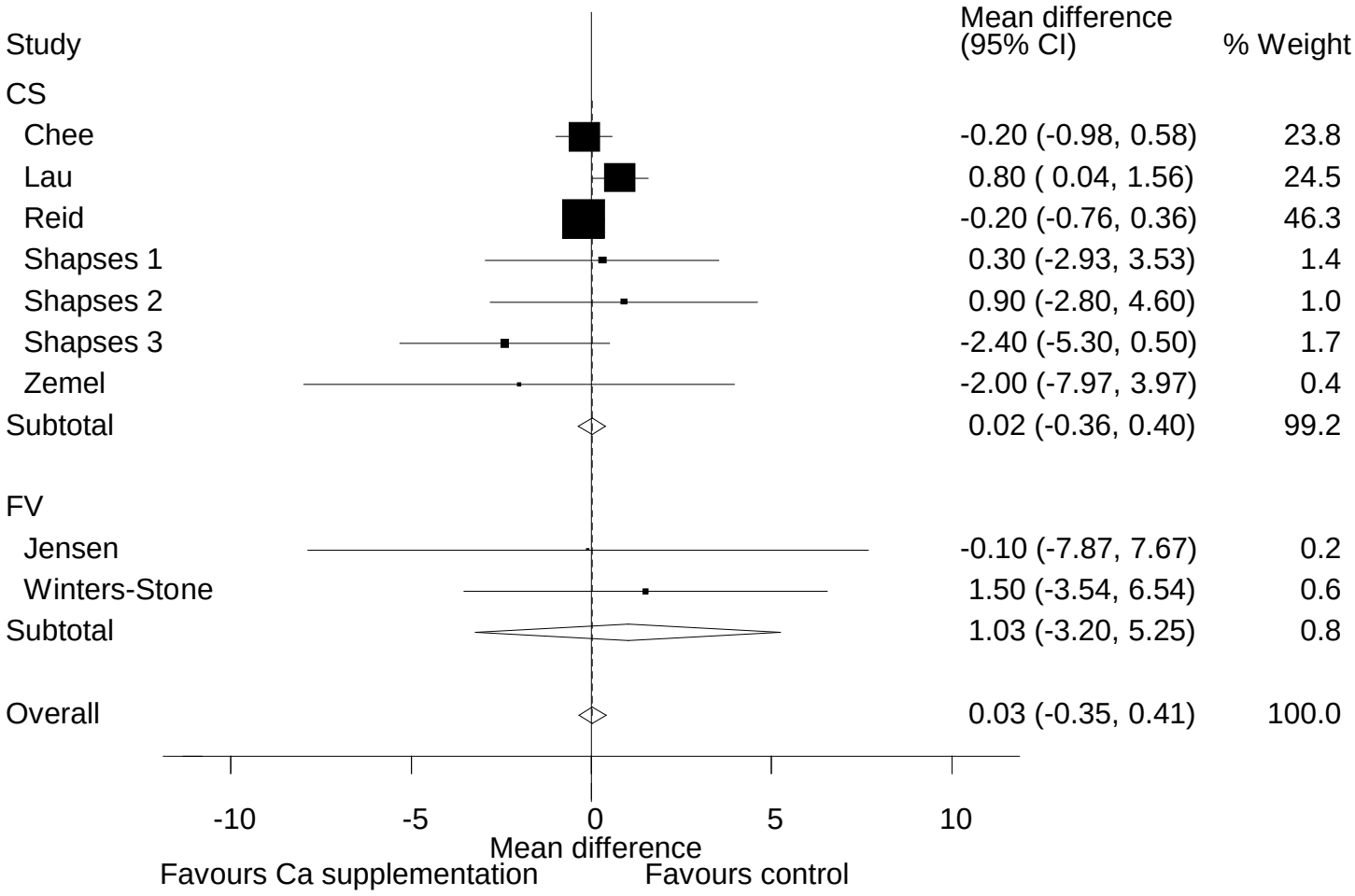
# Option 4: Meta-analysis using a mix of results from different analysis methods

# Calcium supplementation on body weight (Trowman 2006 Br J Nut)

| Trial | Year | Baseline (weight kg) | | | | Follow-up (weight kg) | | Change (weight kg) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Intervention | | Control | | Intervention | Control | Intervention | Control |
| | | N | Mean (SD) | n | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) |
| Chee | 2003 | 91 | 56.1 (8.9) | 82 | 57.2 (9.4) | | | **0.0 (2.6)[a]** | **0.2 (2.6)[a]** |
| Jensen | 2001 | 25 | 94.6 (14.0)[a] | 27 | 93.8 (14.0)[a] | **89.0 (12.7)[a]** | **89.1 (14.7)[a]** | | |
| Lau | 2001 | 95 | 56.9 (7.1) | 90 | 58.9 (7.5) | | | **0.5 (2.6)[a]** | **-0.3 (2.7)[a]** |
| Reid | 2002 | 111 | 66.0 (10.0) | 112 | 68.0 (11.0) | | | **-0.3 (1.8)** | **-0.1 (2.4)** |
| Shapses 1[c] | 2004 | 17 | 84.1 (9.4) | 19 | 89.4 (10.3) | | | **-7.0 (4.6)** | **-7.3 (5.3)** |
| Shapses 2[c] | 2004 | 11 | 85.9 (9.2) | 11 | 94.2 (15.7) | | | **-6.7 (2.6)** | **-7.6 (5.7)** |
| Shapses 3[c] | 2004 | 18 | 93.7 (13.6) | 24 | 93.5 (14.3) | | | **-6.7 (5.5)** | **-4.3 (3.5)** |
| Winters-Stone | 2004 | 13 | 57.2 (4.9) | 10 | 54.1 (7.2) | **56.3 (4.3)** | **54.8 (7.2)** | | |
| Zemel | 2004 | 11 | 99.8 (14.9) | 10 | 103.1 (19.3) | | | **-8.6 (5.3)[a]** | **-6.6 (8.2)[a]** |

a Calculated from the standard error
b Follow-up sample size ntrt = 24 and nctrl = 24
c Shapses *et al* (Shapses *et al*, 2004) report on three randomised controlled trials.
Trials 1, 2, and 3 include postmenopausal women, postmenopausal women special diet, and premenopausal women respectively.

# Combining intervention estimates from SAFV and SACS



# Which option?

| | | Meta-analysis options | | |
|---|---|---|---|---|
| **Domain** | | **2) Only ANCOVA** | **3) Only one (SAFV or SACS)** | **4) Mix (SAFV, SACS, and ANCOVA)** |
| Bias (few small trials) | Chance baseline imbalance across trials | ✓ | | |
| | Selective reporting | ✓ | ✓ | |
| Precision | No heterogeneity | ✓ | | |
| | Heterogeneity | = | = | = |
| Practical issues | | ✗ | ✗ | ✓ |

***McKenzie 2015 Res Synth Methods***

# Which option?

- In many circumstances combining estimates calculated from a mix of analysis methods is reasonable (option 4)
  - Include available ANCOVA estimates where possible
- If combining estimates based on only one analytical method (SAFV or SACS), options for choice of method include
  - the analysis method most frequently reported
  - the analysis method that is likely to yield the greatest precision (correlations likely to be large → use SACS; correlations likely to be small → use SAFV)

# References


- Balk EM, Earley A, Patel K, Trikalinos TA, Dahabreh IJ: Empirical Assessment of Within-Arm Correlation Imputation in Trials of Continuous Outcomes. Methods Research Report. (Prepared by the Tufts Evidence-based Practice Center under Contract No. 290-2007-10055-I.). In: AHRQ Publication. Rockville, MD: Agency for Healthcare Research and Quality; 2012.
- McKenzie JE, Herbison GP, Deeks JJ. Impact of analysing continuous outcomes using final values, change scores and analysis of covariance on the performance of meta-analytic methods: a simulation study. Res Synth Methods. 2015.
- Trowman R, Dumville JC, Hahn S, Torgerson DJ: A systematic review of the effects of calcium supplementation on body weight. Br J Nut 2006, 95(6):1033-1038.
- Wiebe N, Vandermeer B, Platt RW, Klassen TP, Moher D, Barrowman NJ: A systematic review identifies a lack of standardization in methods for handling missing variance data. J Clin Epi. 2006, 59(4):342-353.

# Meta-analysis of skewed data

# Meta-analysis of skewed data

- Standard meta-analytic methods assume normality in the distribution of the means (not raw data)
- Many outcomes are not normally distributed. Examples include:
  - Concentrations
    - e.g. urinary iodine
  - Ratio or reciprocal measures
    - e.g. ratio of partial pressure of arterial oxygen to fraction of inspired oxygen
  - Resource use
    - e.g. length of stay
  - Assessment scales
    - e.g. large proportion of ‘normal’ participants fall towards one extreme of the scale

***Higgins 2008 Stats in Med***

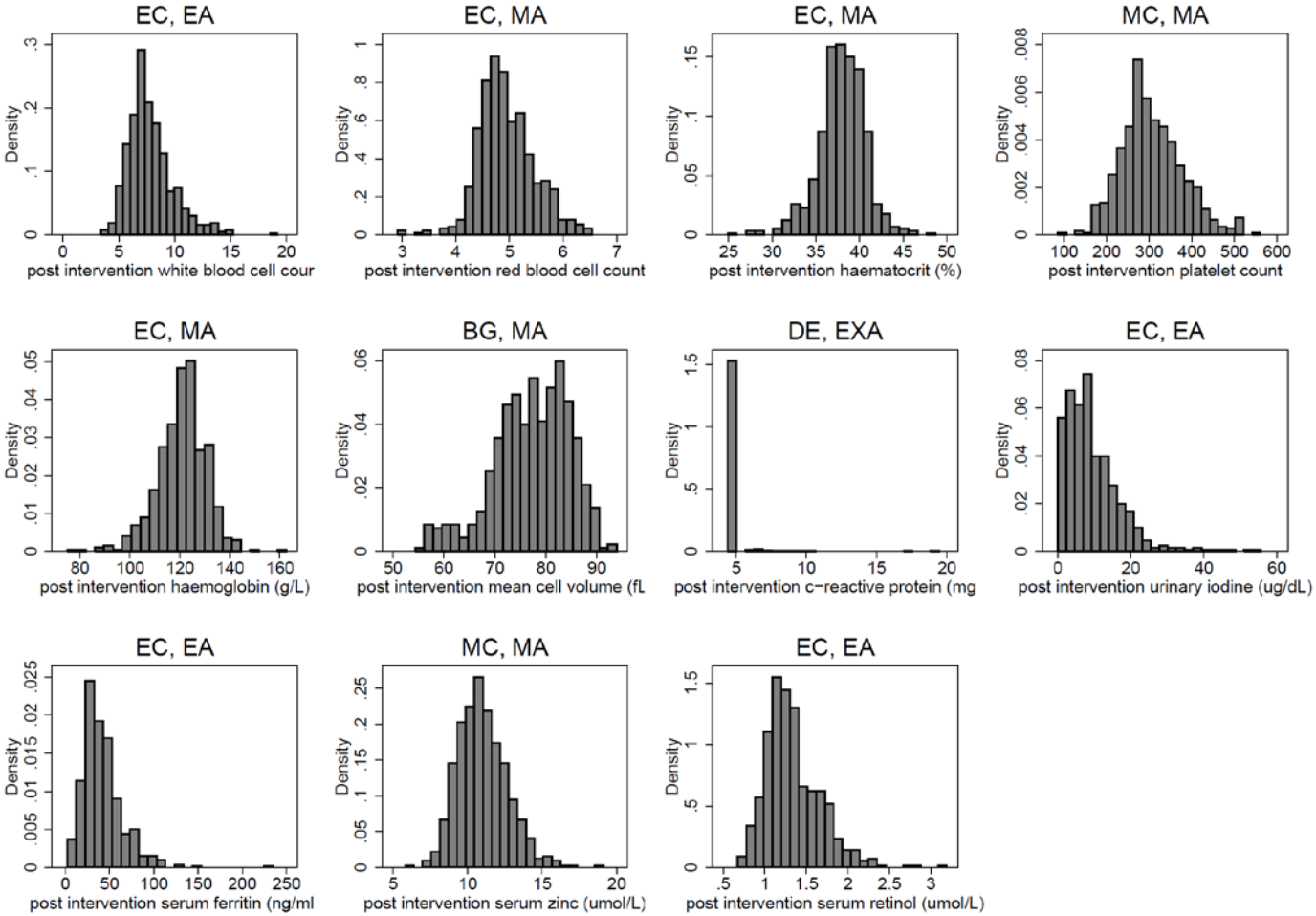
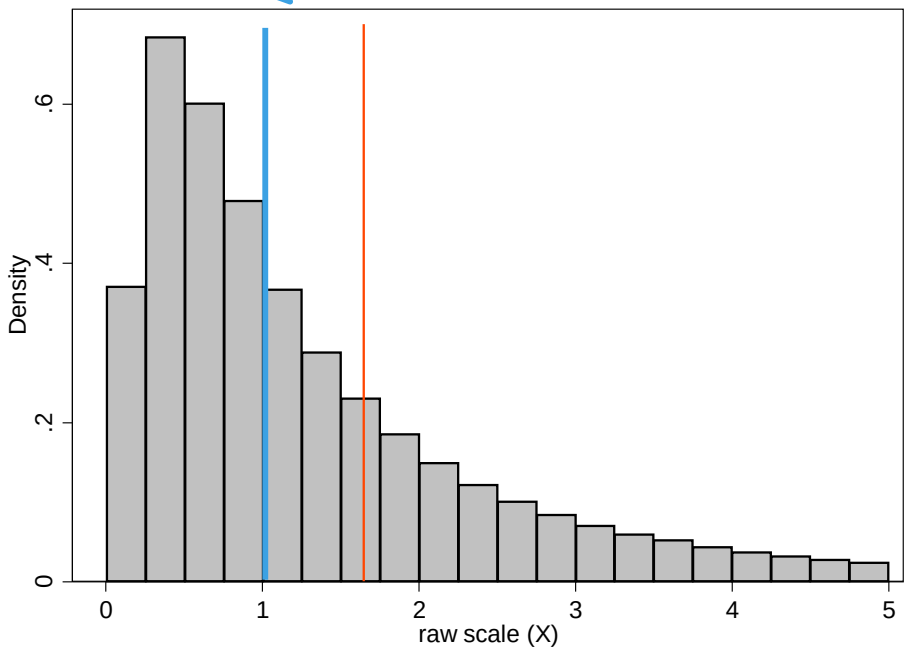



Figure 3: Histograms of post biochemical variables from Thailand RCT

First abbreviation = classification based on tail weight (U = Uniform, BG = Below Gaussian, G = Gaussian, MC = Moderate Contamination, EC = Extreme Contamination, DE = Double Exponential), second abbreviation = classification based on asymmetry (S = Near Symmetry, MA = Moderate Asymmetry, EA = Extreme Asymmetry, EXA = Exponential Asymmetry).

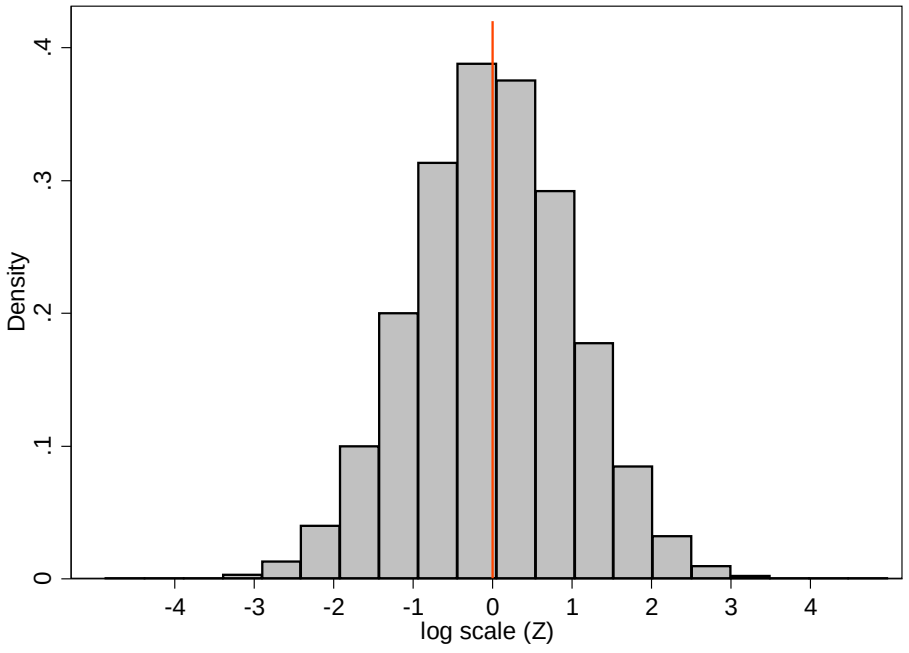
# Dealing with skewed data in studies

Common approach to dealing with skewed data is to log transform the observations, then undertake the analysis on the log scale




$G$ $\bar{X}$

$\mathrm{SD_X}$




$\bar{Z}$

$\mathrm{SD_Z}$

# Meta-analysis of skewed data

- Meta-analytic methods are likely to be valid in large trials even when the outcome is skewed, but practical issues arise
- Same outcome: different scales (raw, log), different statistics


Study 1: Mean, SD, raw scale $(\bar{X}, SD_X)$

Study 2: Mean, SE, raw scale $(\bar{X}, SD_X/n)$

Study 3: Mean, SD, log scale $(\bar{Z}, SD_Z)$

Study 4: Geometric mean, CI $(G, G_l, G_U)$

…

Study X

# Meta-analysis of skewed data

- Aim to include as many trials in the one meta-analysis as possible
- To achieve this, we need to transform the summary statistics from one scale to another

STATISTICS IN MEDICINE
*Statist. Med.* 2008; **27**:6072–6092
Published online 17 September 2008 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.3427

## Meta-analysis of skewed data: Combining results reported on log-transformed or raw scales

Julian P. T. Higgins*, †, Ian R. White and Judith Anzures-Cabrera

*MRC Biostatistics Unit, Institute of Public Health, Robinson Way, Cambridge CB2 0SR, U.K.*

# Scenario 1:
# Convert raw scale to log scale

**Method 1:** Transform $\bar{X}_i$ and $SD_{X_i}$ within each group

**Method 2:** Transform $\bar{X}_i$ and $SD_{X_i}$ assuming a common underlying SD on the log scale

Methods assume X follows a log normal distribution (i.e. $Z = \ln(X) \sim N(\mu, \sigma_Z^2)$) and utilise standard transformation

Two methods *(ad hoc, Taylor series)* for calculating $SE(D_Z)$

Trials requiring conversion

$\bar{X}_i, SD_{X_i}$

$\bar{Z}_i', SD_{Z_i}'$

**Standard result:**

If variable *X* follows a log normal distribution, then:

$$E(X) = exp\left(\mu + \frac{\sigma_Z^2}{2}\right)$$

$$var(X) = (exp(\sigma_Z^2) - 1)exp(2\mu + \sigma_Z^2)$$

# Scenario 1: Convert raw scale to log scale

**Method 1:** Transform $\bar{X}_i$ and $SD_{X_i}$ within each group

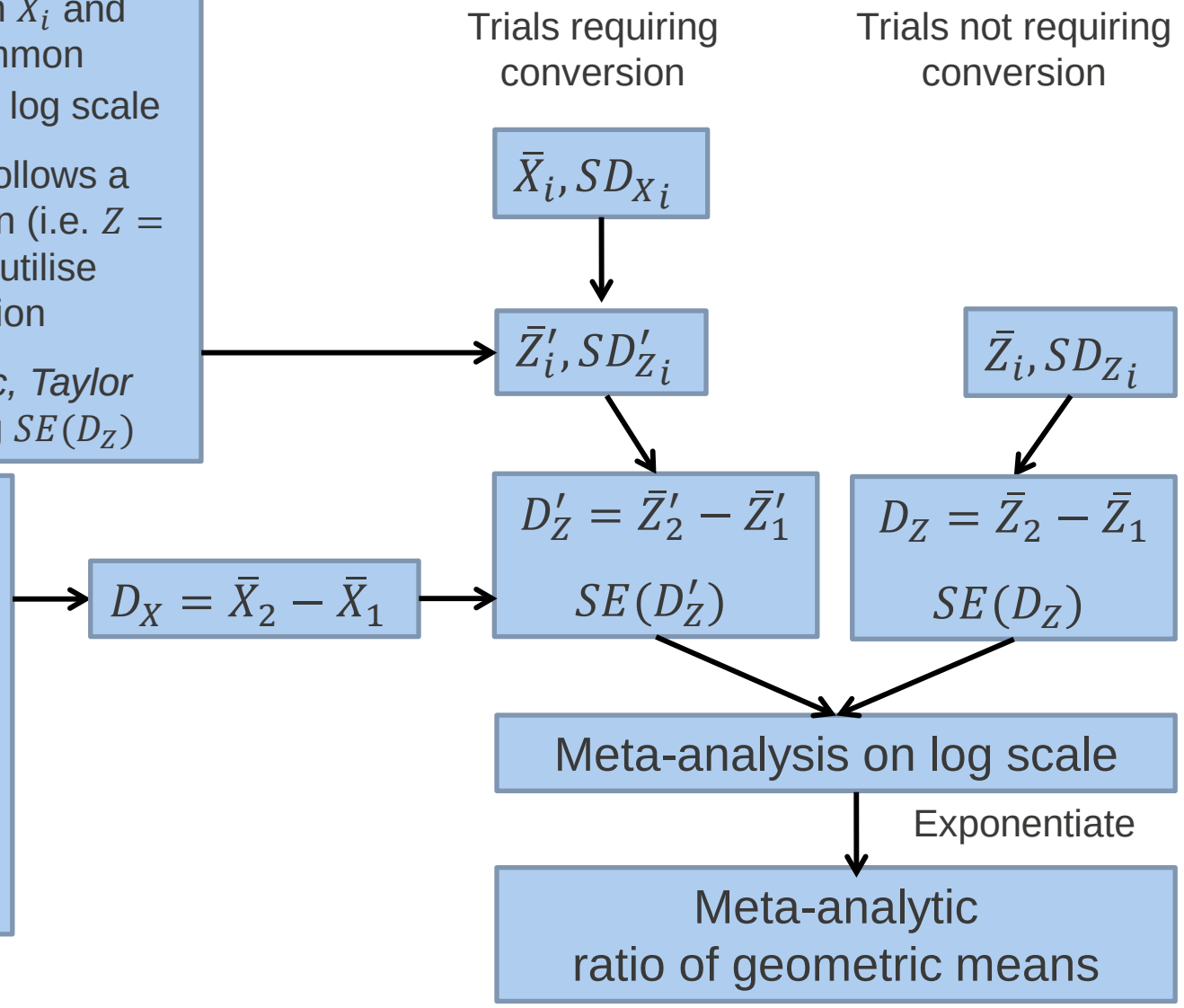
**Method 2:** Transform $\bar{X}_i$ and $SD_{X_i}$ assuming a common underlying SD on the log scale

Methods assume X follows a log normal distribution (i.e. $Z = \ln(X) \sim N(\mu, \sigma_Z^2)$) and utilise standard transformation

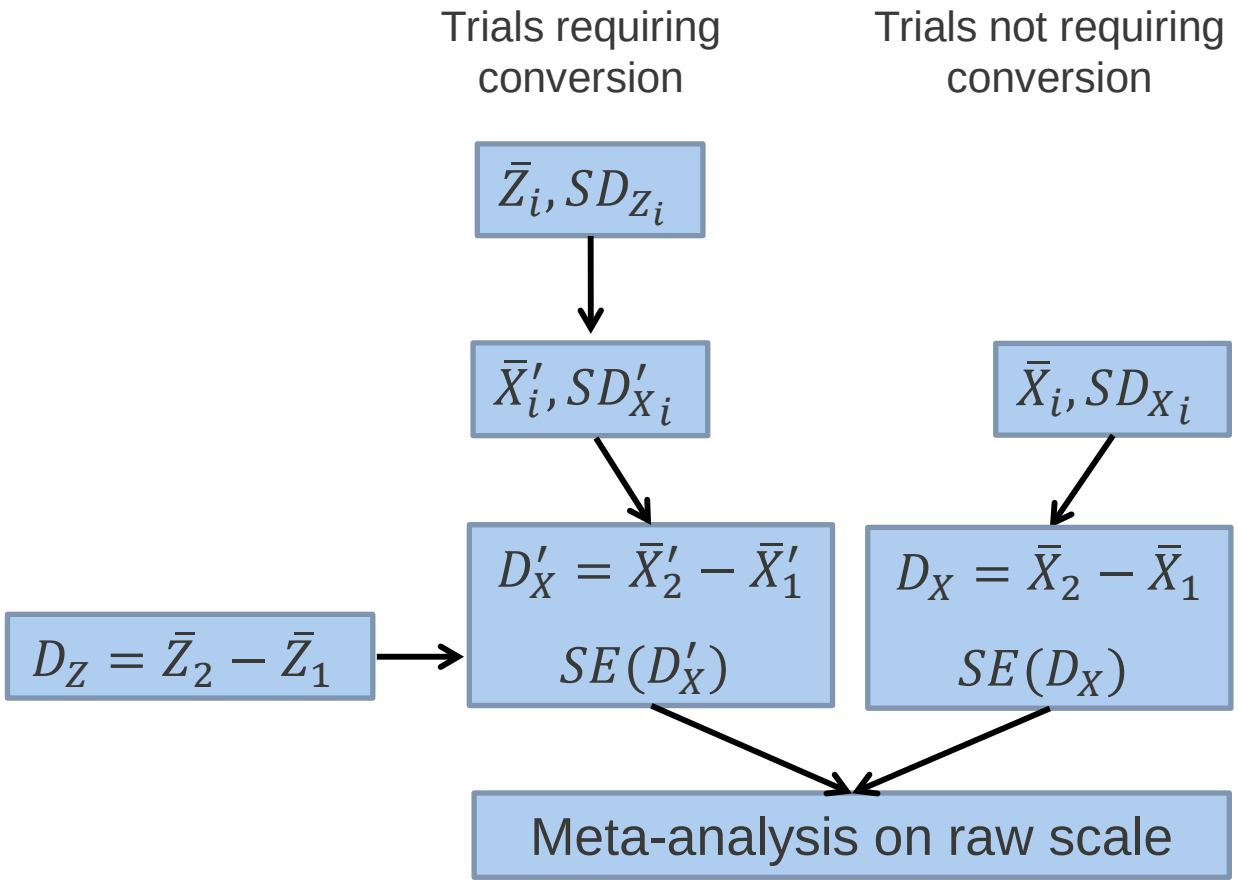
Two methods *(ad hoc, Taylor series)* for calculating $SE(D_Z)$

**Method 3:** Targets difference between the groups ($D_X$), rather than group means.

Method does *not* assume X follows a log normal distribution



# Scenario 2:
# Convert log scale to raw scale



## Meta-analysis from a systematic review (Sagoo et al)

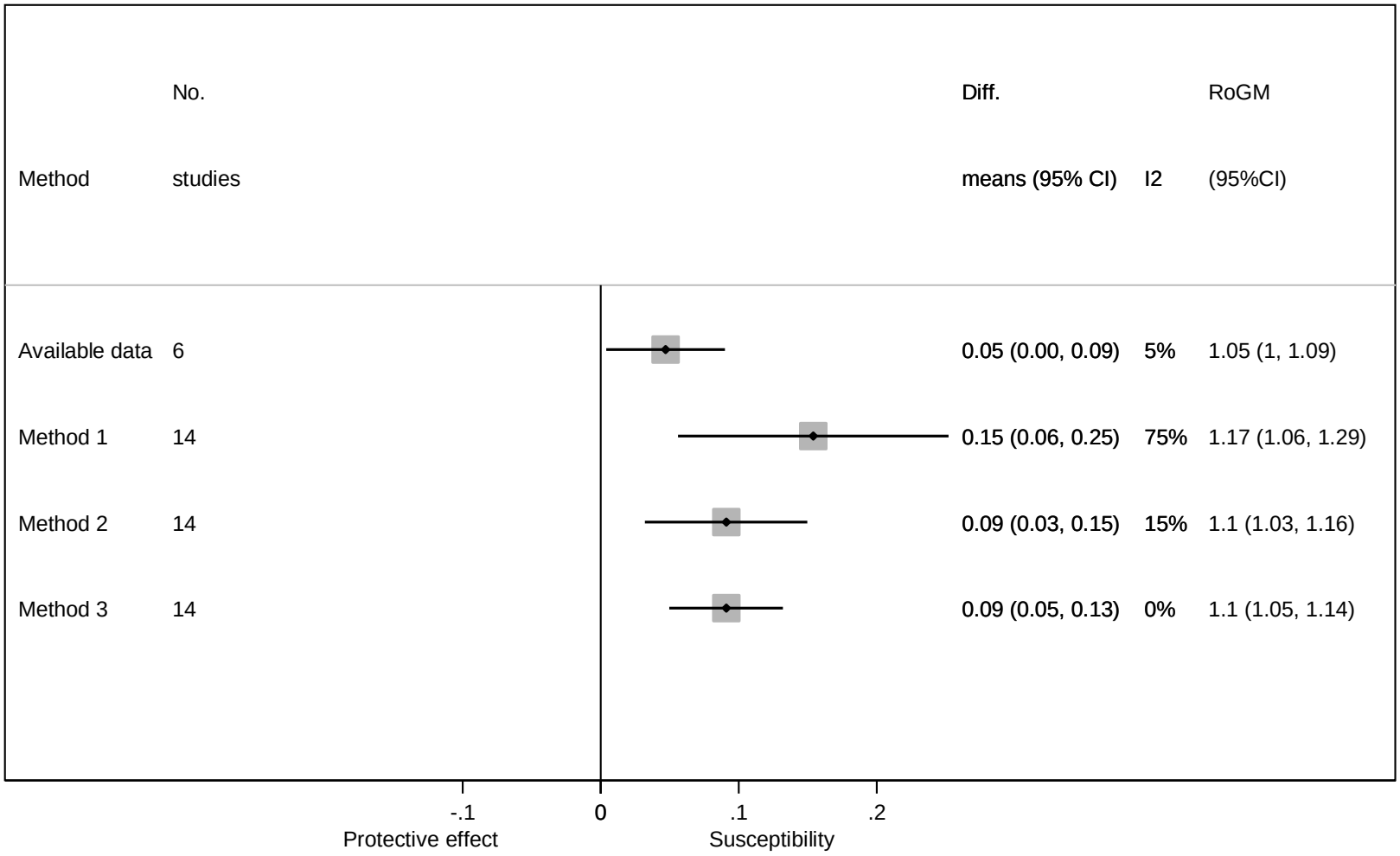
Examines association between triglyceride level and being a carrier or non-carrier of the D9N polymorphism in the LPL gene.

Table I. Data available for D9N polymorphism in the lipoprotein lipase gene and triglyceride levels.

| | Carriers | | | | | Non-carriers | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Raw | | Log | | | Raw | | Log | |
| | *n* | Mean | SD | Mean | SD | *n* | Mean | SD | Mean | SD |
| Boer 2003b | 34 | — | — | 0.31 | 0.58 | 1002 | — | — | 0.33 | 0.53 |
| Copenhagen | 241 | 2.10 | 1.46 | 1.05 | 0.37 | 8429 | 1.85 | 1.54 | 0.98 | 0.34 |
| CDRFMP | 14 | 2.05 | 1.21 | — | — | 364 | 1.57 | 1.11 | — | — |
| EARS I & II | 71 | 1.12 | 0.34 | — | — | 1608 | 0.99 | 0.80 | — | — |
| ECTIM | 22 | 1.82 | 1.46 | — | — | 784 | 1.84 | 1.47 | — | — |
| Ehrenborg 1997 | 15 | 1.01 | 0.36 | — | — | 77 | 0.99 | 0.53 | — | — |
| Ferencak 2003 | 5 | 2.04 | 0.92 | — | — | 195 | 1.81 | 0.84 | — | — |
| FOS | 58 | 1.61 | 0.72 | — | — | 2200 | 1.38 | 1.16 | — | — |
| Glisic 2003b | 4 | 2.42 | 1.53 | 0.74 | 0.60 | 129 | 1.64 | 0.94 | 0.37 | 0.49 |
| Reykjavik | 10 | 1.64 | 1.64 | 0.20 | 0.74 | 274 | 1.04 | 0.49 | −0.05 | 0.42 |
| Rios 2003 | 10 | 1.60 | 0.70 | 0.39 | 0.41 | 187 | 1.75 | 0.92 | 0.43 | 0.50 |
| Schulte 1996 | 17 | 1.96 | 0.82 | — | — | 644 | 1.56 | 0.82 | — | — |
| Talmud 1998 | 12 | 1.35 | 0.52 | — | — | 96 | 1.27 | 0.52 | — | — |
| Yang 2004 | 235 | 2.39 | 1.46 | 0.74 | 0.50 | 1275 | 2.34 | 1.26 | 0.73 | 0.49 |

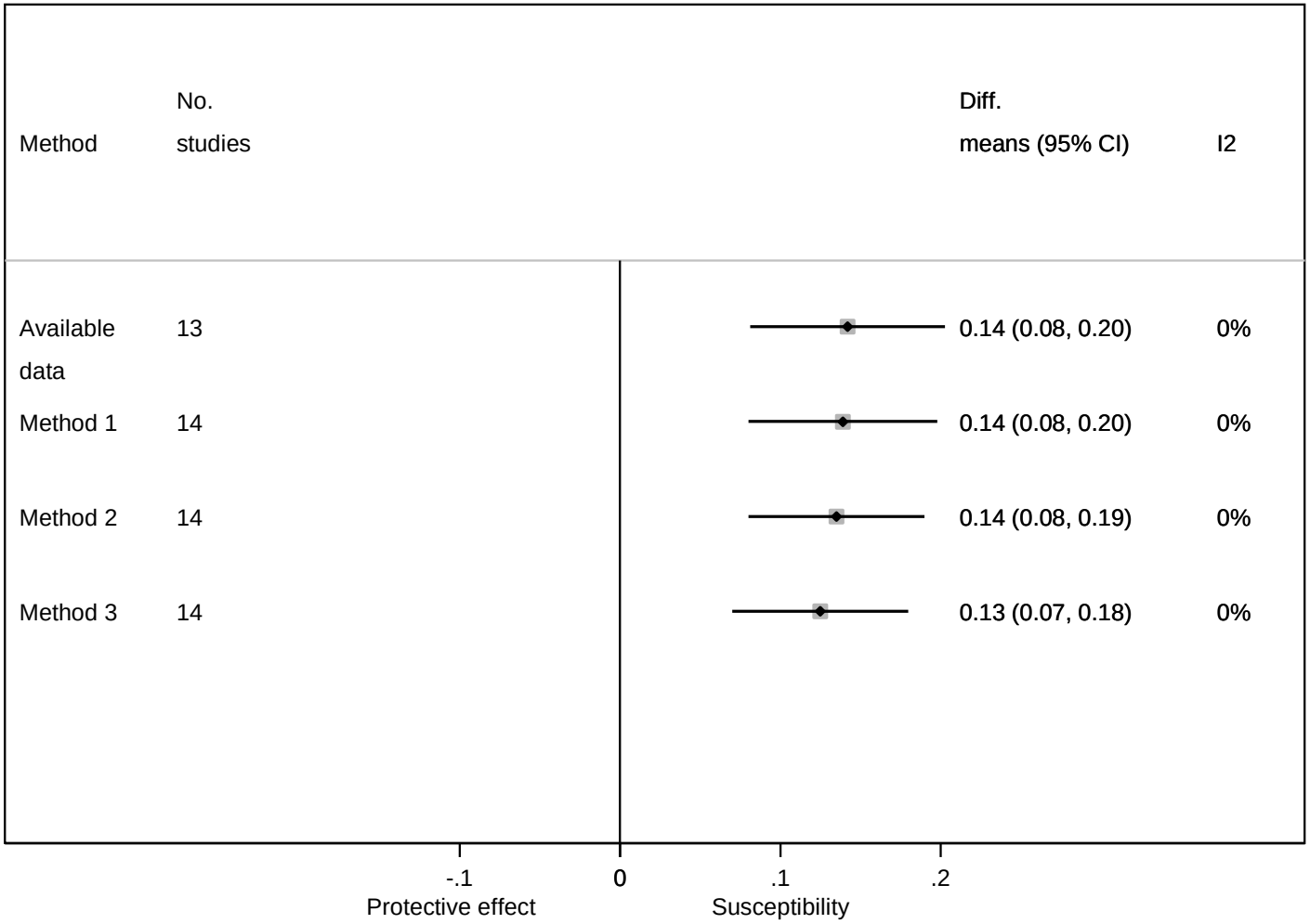
# Raw to log scale

Difference in mean log triglyceride level (Carriers – Non-carriers) using alternative conversions



# Log to raw scale

Difference in mean raw triglyceride level (Carriers – Non-carriers) using alternative conversions



# Which conversion method? Which direction?

*Which conversion method?*

- Not a uniformly preferable method
- All methods reasonably robust to data having distributions other than log normal
- Method 1 preferable when SDs differ across groups; when SDs are similar, greater precision obtained using Method 2

*Which direction? Raw to log scale OR log to raw scale?*

- Use the scale most frequently used
- Scale best meeting meta-analytic assumptions (scale believed to be less skewed)
- Meta-analysis on the log-scale may reduce heterogeneity

# Length of intubation (hours)

| Study | High dose opioid | | | | | Low dose opioid | | | | | Significance as reported by authors |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | Mean | SD | Median | Range | n | Mean | SD | Median | Range | |
| Slogoff 1989<br>Enflurane<br>Halothane<br>Isoflurance | 254 | 22.8 | 12.3 | | | 257<br>253<br>248 | 14.5<br>16.8<br>14.7 | 6.3<br>7.2<br>5.4 | | | ANOVA<br>p = 0.001 |
| Bell 1994 | 19 | | | 12.96 | | 20 | | | 4.42 | | p = 0.0005 |
| Cheng 1996 | 51 | 18.9 | 1.4 | | | 51 | 4.1 | 1.1 | | | p < 0.02 |
| Myles 1997 | 66 | 21.5 | 5.1 | 12.3 | 3.5-31.5 | 58 | 11.4 | 9.9 | | | p = 0.006 |
| Silbert 1998 | 42 | | | 7.0 | 2.1-19 | 38 | | | 4.0 | 0.5-15.5 | p < 0.01 |
| Michalopoulos 1998 | 72 | 11.6 | 1.3 | | | 72 | 7.3 | 0.7 | | | p = 0.0001 |
| Sakaida 1998 | 20 | 14.5 | 4.5 | 15 | 6-25.3 | 20 | 5.6 | 1.6 | 5.2 | 3.5-9 | p < 0.05 |
| Berry 1998 | 42 | | | 12.62 | 8.32-20.67 | 43 | | | 1.83 | 0.1-4.25 | significant |
| Myles 2001 | 24 | | | 9.7 | 1.1-25 | 24 | | | 6.5 | 0.4-150 | ?significant |

Anesthesiology 2003; 99:982–7



***A Systematic Review of the Safety and Effectiveness of Fast-track Cardiac Anesthesia***

*Paul S. Myles, M.B.B.S., M.P.H., M.D., F.C.A.R.C.S.I., F.A.N.Z.C.A.,* David J. Daly, M.B.B.S., F.A.N.Z.C.A.,†*
*George Djaiani, M.D., D.E.A.A., F.R.C.A.,‡ Anna Lee, B.Pharm., M.P.H., Ph.D.,§*
*Davy C. H. Cheng, M.D., M.Sc., F.R.C.P.C.‖*

Source: Doug Altman

# Dealing with non-parametric statistics

- Assume that the data are (close to) lognormal
  - median ~ geometric mean
  - Convert centiles to mean and SD on the log scale
    - Converting ranges see: Walter 2007 J Clin Epi; Hozo 2005 BMC Med Res Methodol
    - Converting IQR: Section 7.7.3.5 Cochrane Handbook; Wan 2014 BMC Med Res Methodol

- Combine using conversion approaches in Higgins 2008 Stats Med

# References


- Higgins JPT, Deeks JJ (editors). Chapter 7: Selecting studies and collecting data. In: Higgins JPT, Green S (editors), Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
- Higgins JP, White IR, Anzures-Cabrera J: Meta-analysis of skewed data: combining results reported on log-transformed or raw scales. Stat Med 2008, 27(29):6072-6092.
- Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol. 2005; 5: 13.
- Walter SD, Yao X: Effect sizes can be calculated for studies reporting ranges for outcome variables in systematic reviews. Journal of clinical epidemiology 2007, 60(8):849-852.
- Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Medical Research Methodology. 2014; **14**: 135.