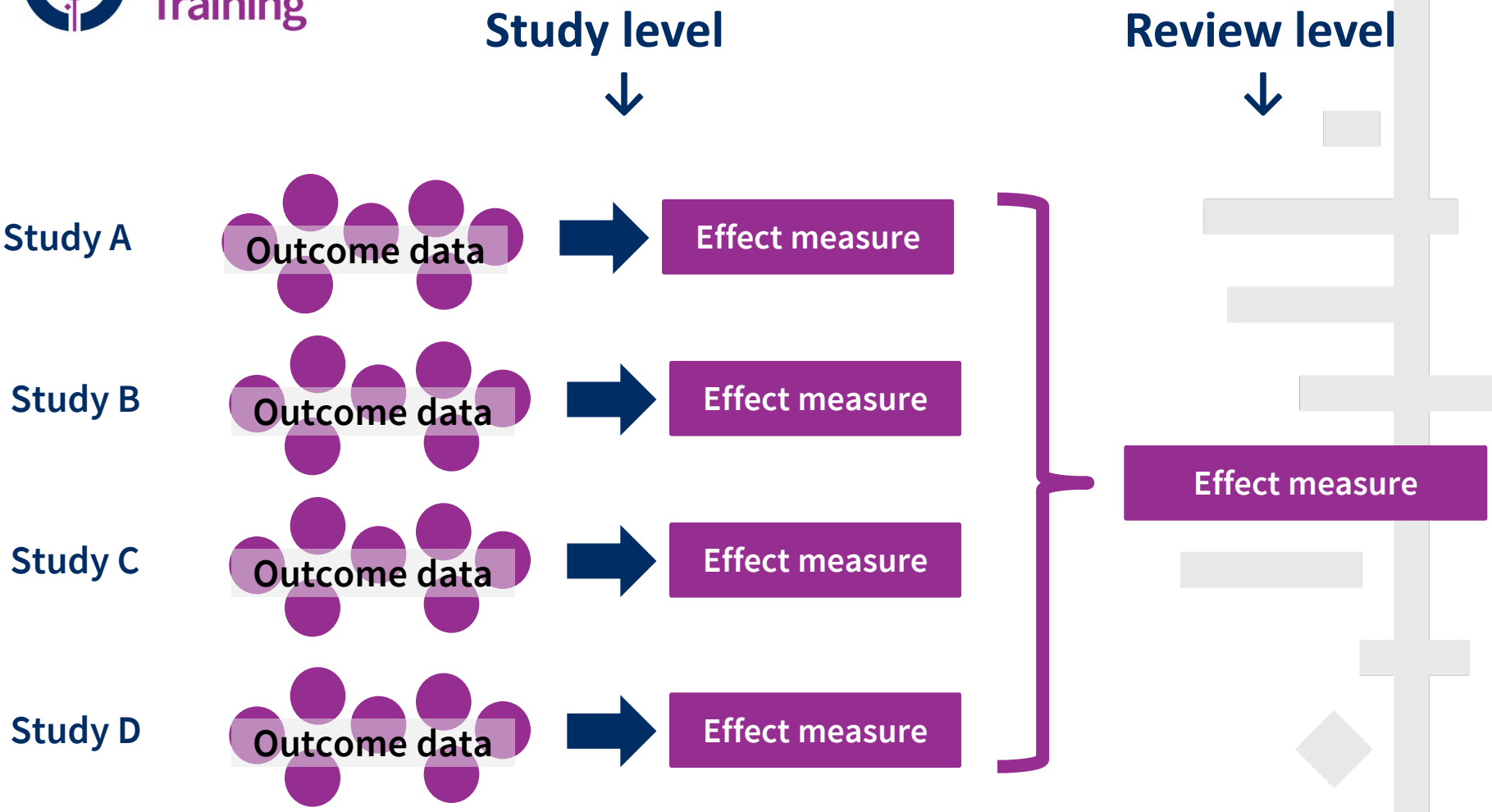
Cochrane Training

Dichotomous and Continuous Outcomes

Trusted evidence.
Informed decisions.
Better health.

Cochrane Training



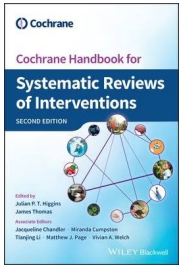
Source: Jo McKenzie & Miranda Cumpston

Types of outcome data

- **Dichotomous:** Two categories (event or no event)
  - Alive or dead
  - Healed or not healed
  - Pregnant or not pregnant

- **Continuous:** measured on a scale (range of values)
  - Height, weight
  - Other numerical scales commonly assessed as continuous such as quality of life, pain, depression

Part 1:
Dichotomous outcomes



See Chapters 6 & 10 of the Handbook

What are dichotomous outcomes?

- **Definition:** when the outcome for every participant is one of two possibilities or events
- We express the **chance** of the event using **Risk** or **Odds**
- Although risk and odds are both used for expressing the chance of being in one of two groups, they are calculated differently

Calculating Risk

- 24 people drank coffee
  6 developed a headache
- Risk of a headache

  = 6 people with headache / 24 people who could have had one

  = 6/24 = ¼ = 0.25 = 25%

$$\text{Risk} = \frac{\text{no. participants with event of interest}}{\text{total no. participants}}$$

Cochrane Training

Calculating Odds

- 24 people drank coffee
  6 developed a headache
- Odds of a headache

  = 6 people with headache/18 without headache

  = 6/18 = 1/3 = 0.33 = 1:3 (not usually as %)

$$\text{Odds} = \frac{\text{no. participants with event of interest}}{\text{no. participants without event of interest}}$$

Do risks and odds differ much?

Two examples from caffeine trials

- 5 people with ‘headache’ out of 65
- Chance of having a headache

Risk: 5/65 = 0.077 **Odds:** 5/60 = 0.083

Rare event

- 130 people ‘still awake’ out of 165
- Chance of still being awake

Risk: 130/165 = 0.79 **Odds:** 130/35 = 3.71

Common event

Comparing two groups: the 2x2 table

	Event	No event	Total
Intervention	a	b	**a + b**
Control	c	d	**c + d**
Total	**a + c**	**d + b**	**a + b + c + d**

Comparing two groups: the 2x2 table

	Event: Headache	**No event:** No headache	**Total**
Intervention: Caffeine	17	51	**68**
Control: Decaf	9	55	**64**
Total	**26**	**106**	**132**

Cochrane Training

Comparing two groups: effect measures

- Effect measures for dichotomous outcomes
  - Risk Ratio (**RR**) (relative risk)
  - Odds Ratio (**OR**)
  - Risk Difference (**RD**) (absolute risk reduction)
- All estimates are uncertain, and should be presented with a confidence interval

Risk ratio

- **Formula:** $\text{Risk ratio} = \dfrac{\text{Intervention risk}}{\text{Control risk}}$

- Risk of event with intervention: **17/68**
- Risk of event with control: **9/64**

	Headache	No headache	Total
Intervention: Caffeine	17	51	68
Control: Decaf	9	55	64
Total	**26**	**106**	**132**

$$\textbf{Risk ratio} = \frac{17/68}{9/64} = \frac{0.25}{0.14} = \textbf{1.79}$$

Risk ratio interpretation

- **RR = 1:** there is no difference between the groups
- **RR > 1:** increased risk in the intervention group
  - For our calculated RR of 1.79, this means that the intervention increased the risk of headache by 79% [100*(RR-1)%]
  - Or the risk of having a headache in the intervention group was 179% of the risk in the control group
- **RR < 1:** decreased risk in the intervention group
  - For example, an RR of 0.79 means that the intervention reduced the risk of headache by 21% [100*(1-RR)%]
  - Or the risk of having a headache with treatment was 79% of the risk in the control group

Odds ratio

- **Formula:** $\text{Odds ratio} = \dfrac{\text{Intervention odds}}{\text{Control odds}}$

- Odds of event with intervention: **17/51**
- Odds of event with control: **9/55**

	Headache	No headache	Total
Intervention: Caffeine	17	51	68
Control: Decaf	9	55	64
Total	**26**	**106**	**132**

$$\text{Odds ratio} = \frac{17/51}{9/55} = \frac{0.33}{0.16} = 2.06$$

Odds ratio interpretation

- **OR = 1:** there is no difference between the groups
- **OR > 1:** increased odds in the intervention group
  - For our calculated OR of 2.06, this means that the intervention increased the odds of headache by 106% [100*(OR-1)%]
  - Or the intervention increased the odds of headache to 206% of the odds in the control group
- **OR < 1:** decreased odds in the intervention group
  - For example, an OR of 0.06 means that the intervention reduced the odds of headache by 94% [100*(1-OR)%]
  - Or the intervention reduced the odds of headache to 6% of the odds in the control group

Risk difference

- **Formula:** Risk difference = Risk with intervention ĸ Risk with control

- Risk of event with intervention: **17/68**
- Risk of event with control: **9/64**

	Headache	No headache	Total
Intervention: Caffeine	17	51	68
Control: Decaf	9	55	64
Total	**26**	**106**	**132**

Risk di3 erence = 17/68 ĸ 9/64 = 0.25 ĸ 0.14 = 0.11

Cochrane Training

Risk difference interpretation

- **RD = 0:** there is no difference between the groups
- **RD > 0:** increased (absolute) risk in the intervention group
  - For our calculated RD of 0.11, this means that the intervention increased the risk of headache by 11 percentage points
  - Or 14 out of 100 people experienced a headache in the control group. 11 more people experienced a headache in the intervention group
- **RD < 0:** decreased (absolute) risk in the intervention group
  - For example, an RD of -0.11 means that the intervention reduced the risk of headache by 11 percentage points
  - Or 14 out of 100 people experienced a headache in the control group. 11 fewer people experienced a headache in the intervention group

Cochrane Training

Choosing an effect measure

- **Communication of effect**
  - **OR:** hard to understand, often misinterpreted
  - **RR:** easier, but relative (can mean a very big or very small change)
  - **RD:** often easiest (absolute terms; easily converted to natural frequencies or NNT)

- **Consistency of effect**
  - **OR and RR:** less variable across different populations and often more stable across studies with different baseline risks
  - **RD:** can vary more with baseline risk

Source: Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine* 2002; 21:1575-1600

Cochrane Training

Choosing an effect measure

- **Mathematical properties**
  - **RR:** results can be very different depending on how you defined the event of interest (e.g. good or bad, presence of absence)
  - **OR and RD:** result direction often consistent regardless of the definition of the event
  - **OR:** mathematically convenient since it is unbounded

Your Cochrane Group may have a policy

Cochrane Training

Choosing an effect measure: summary

	OR	RR	RD
Communication	✗	✓	✓✓
Consistency	✓	✓	✗
Mathematics	✓✓	✗	✗

Collecting dichotomous outcome data

- Four numbers needed for effect measure and variance:

	Headache	No headache	Total
Caffeine	17	51	68
Decaf	9	55	64
Total	**26**	**106**	**132**

Try to collect the actual number measured for each outcome, at each time point

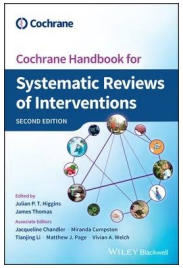
Cochrane Training

Other data formats can also be used

- **Percentages**
  - Number of events can be calculated if sample size is known

- **Overall effect estimate**
  - Where results for each group is not reported separately
  - Can include in meta-analysis using generic inverse variance method
  - Need a measure of variance (e.g. SE, 95% CI)

Part 2:
Continuous outcomes



See Chapters 6 & 10 of the Handbook

What are continuous outcomes?

- **Definition:** measured on a scale
  - Can take any value in a specified range
  - Intervals between values are equally spaced

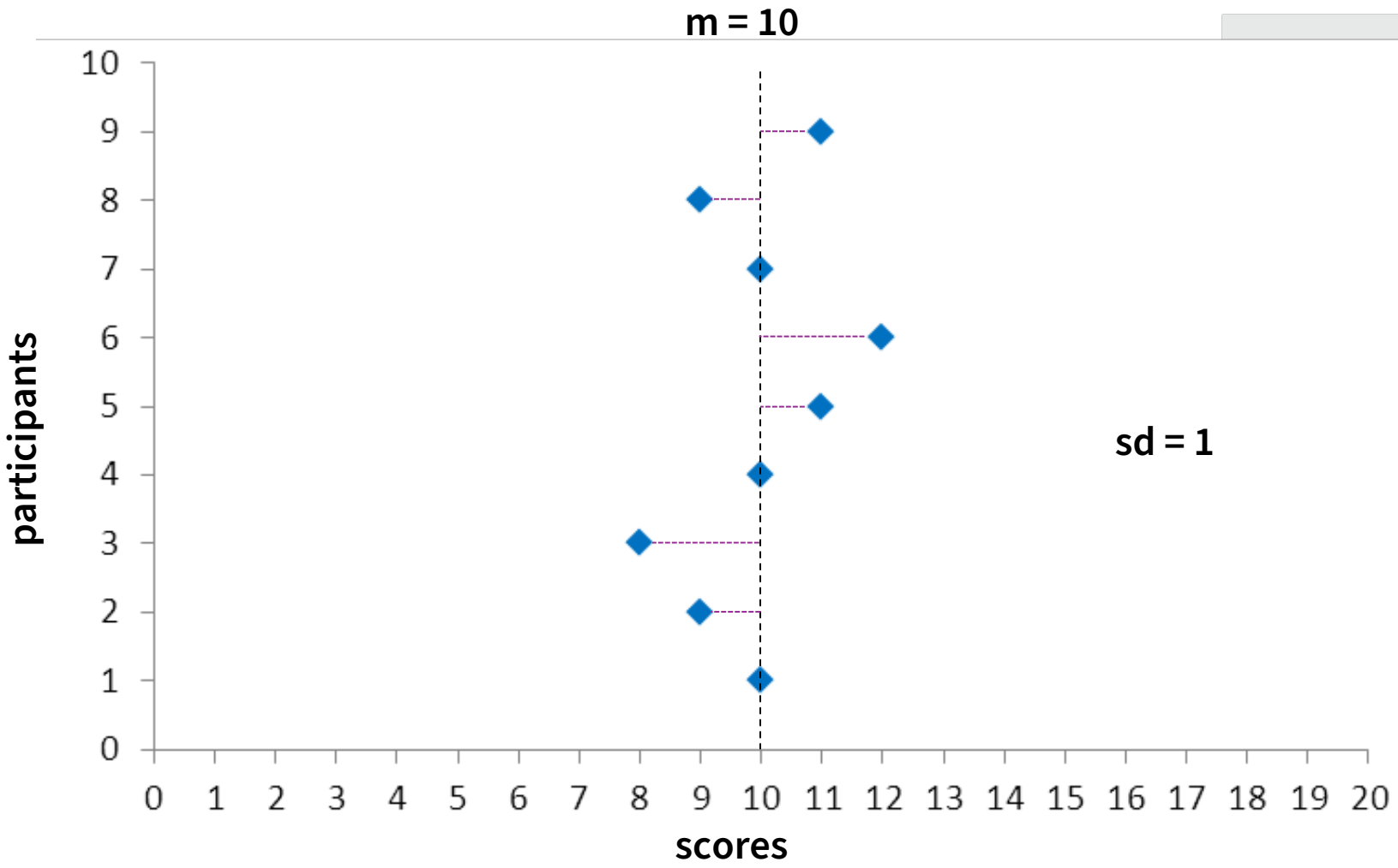
- **Note:** many scales are technically *ordinal* but often treated as continuous in analyses

Expressing continuous outcomes

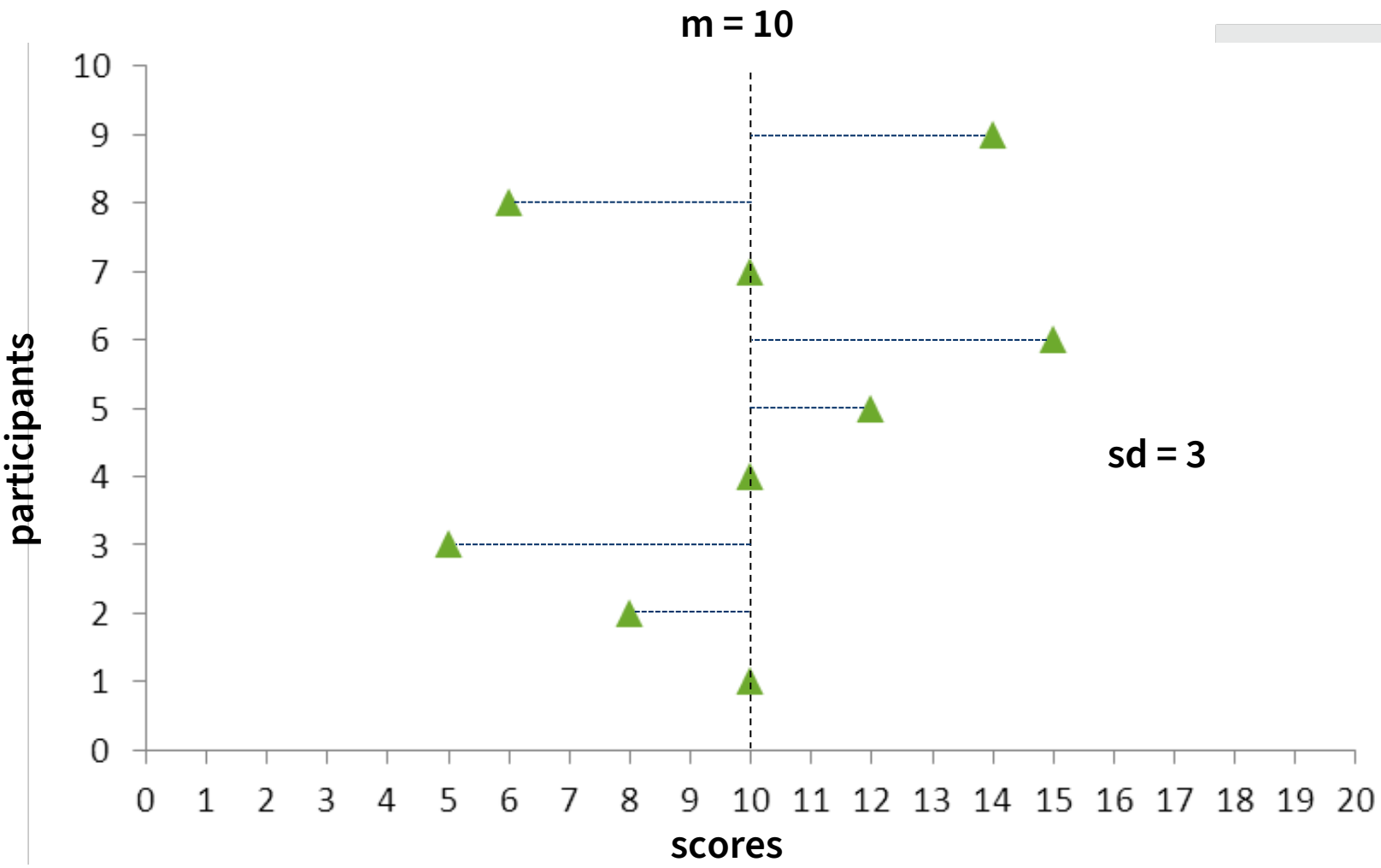
- Two components
  - Mean value
  - Measure of variation

Irritability score	Mean	SD	N
Caffeine	**20**	**9.1**	65
Decaf	**33**	**8.6**	67

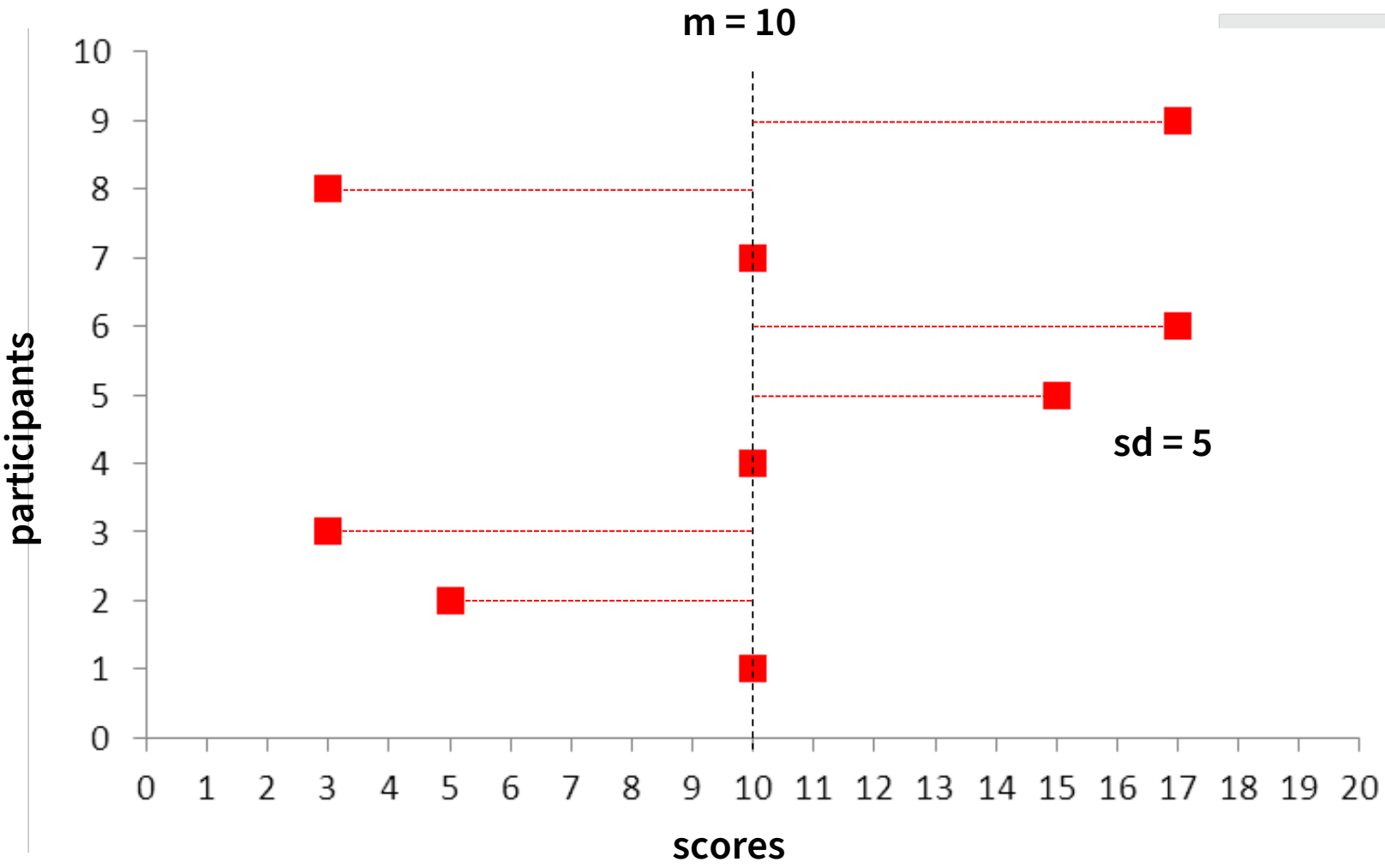
Standard deviations



Standard deviations



Standard deviations



Comparing two groups

- Effect measures for continuous outcomes
  - Mean difference (**MD**) (difference of means)
  - Standardised mean difference (**SMD**)
- All estimates are uncertain, and should be presented with a confidence interval

Mean difference

- When all studies used the same measurement scale
- **Formula:** Mean difference = Mean of intervention − Mean of control

Irritability score	Mean	SD	N
Caffeine	20	9.1	65
Decaf	33	8.6	67

Mean difference = 20 − 33 = −13

Mean difference interpretation

- **MD = 0:** no difference in means
- **MD < 0:** intervention group mean is lower
  - For our calculated MD of -13, this means that on average, participants with the intervention scored 13 points lower on the irritability scale
- **MD > 0:** intervention group mean is higher
  - For example, an MD of 13 means that on average, participants with the intervention scored 13 points higher on the irritability scale

Interpreting mean difference

- **Context is key:** How should we interpret a score of -13?

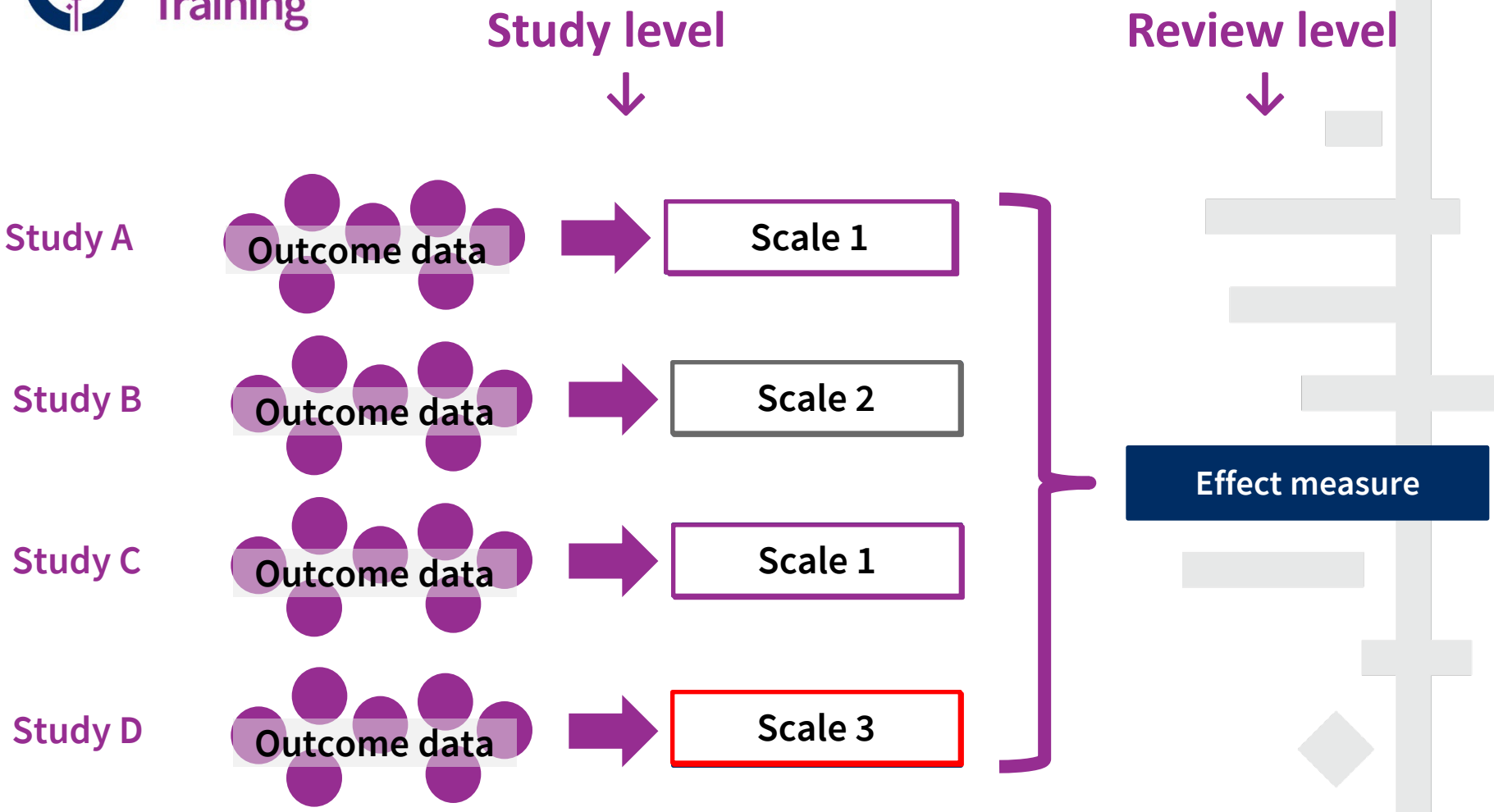
- **Depends on:**
  - Direction of the scale
  - Length of the scale
  - Minimally important difference
  - Good or bad outcome

higher = more irritable

0 - 50

5

bad


Study level
Review level
Study A
Outcome data
Scale 1
Study B
Outcome data
Scale 2
Study C
Outcome data
Scale 1
Study D
Outcome data
Scale 3
Effect measure

Standardised mean difference

- When different scales were used to measure the same outcome
- SMD standardises the results by expressing the difference in *standard deviation units*

- **Formula:** $$\text{SMD} = \frac{\text{Mean of intervention } \kappa \text{ Mean of control}}{\text{Pooled standard deviation of both groups}}$$

- **Caution:** assumes scales measure the same underlying concept

Calculating and interpreting SMD

Irritability score	Mean	SD	N	MD	SMD
Caffeine	**20**	9.1	65	**-13**	**-1.5**
Decaf	**33**	8.6	67		

- **Calculation:** RevMan will run this calculation for you
- **Interpretation:**
  - SMD = -1.5 means that on average, irritability was 1.5 standard deviations lower in the intervention group.
  - Interpretation depends on same factors as mean difference
  - Compare to available SDs
  - Difficult for readers to interpret – convert results to a specific scale for reporting

Converting SMD: example

Review: Manual therapy and exercise for adhesive capsulitis (frozen shoulder)
Comparison: 1 Manual therapy plus exercise plus electrotherapy plus placebo injection versus glucocorticoid injection
Outcome: 1 Overall pain

Study or subgroup	MT+Ex+Electro+Saline Inj N	Mean(SD)	Glucocorticoid injection N	Mean(SD)	Std. Mean Difference IV,Random,95% CI	Weight	Std. Mean Difference IV,Random,95% CI
1 Change from baseline to 6 weeks							
Carette 2003	26	-21.8 (27.0248)	23	-39.1 (26.8567)		51.5 %	0.63 [0.06, 1.21]
Ryans 2005	20	-17.6 (39.1)	17	-9.7 (18.5)		48.5 %	-0.25 [-0.90, 0.40]
Subtotal (95% CI)	**46**		**40**			**100.0 %**	**0.21 [-0.65, 1.07]**

Heterogeneity: Tau² = 0.29; Chi² = 3.93, df = 1 (P = 0.05); I² =75%
Test for overall effect: Z = 0.47 (P = 0.64)

- Pain measured using: SPADI pain scale in Carette 2003 and VAS rest pain in Ryans 2005
- **Convert SMD to SPADI scale:**
  - Multiply SMD and 95% CI by pooled baseline SD (18) from Carette 2003:

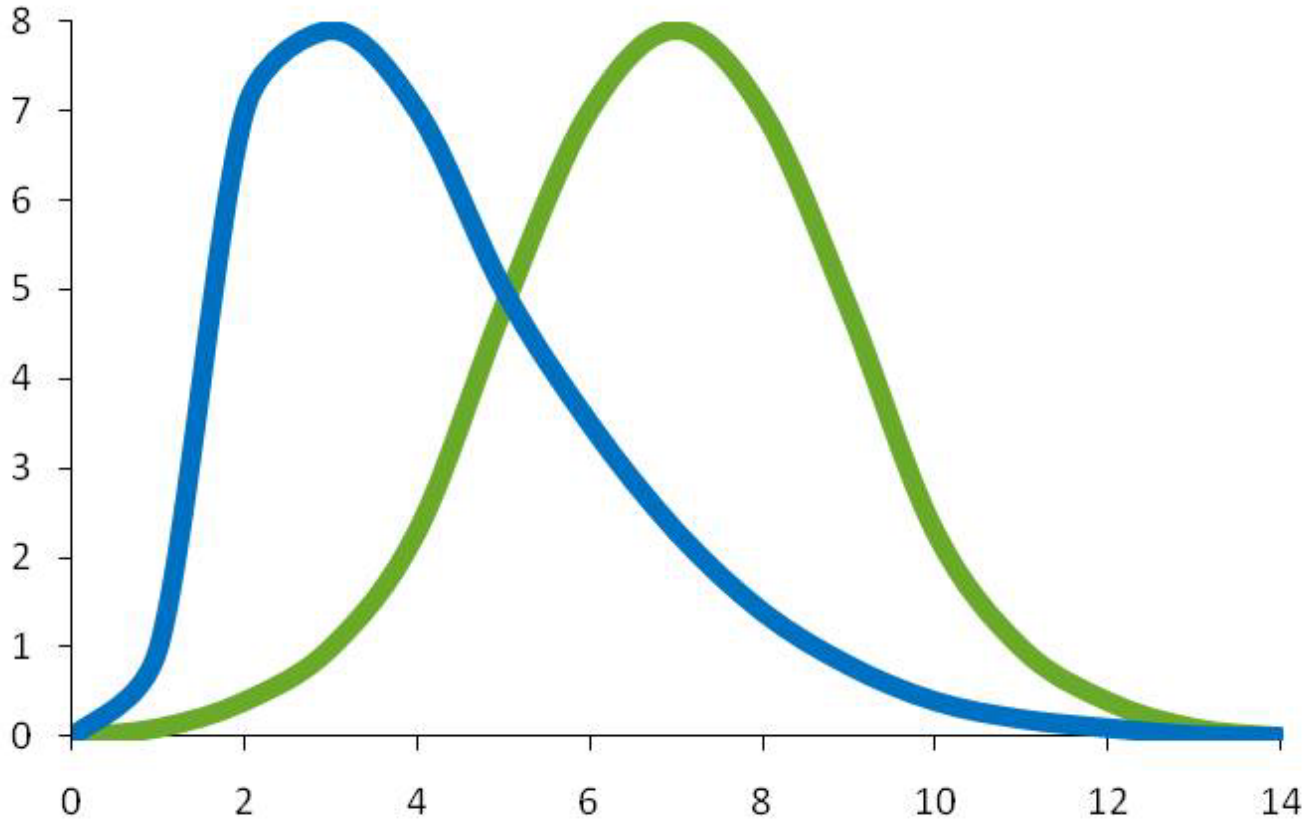
    MD = (SMD x pooled baseline SD) = (0.21 x 18) = 3.78
  - Report result:

    "SMD 0.21, 95% CI -0.65 to 1.07; this is equivalent to a MD of 3.78 points (-11.7 to 19.26) on a 100-point scale (SPADI)"

Source: Page MJ et al. Manual therapy and exercise for adhesive capsulitis (frozen shoulder). Cochrane Database of Systematic Reviews 2014;8:CD011275.

Cochrane Training

Normally distributed data



Skewed data

- **Indications of skew**
  - Reported as geometric mean or median plus interquartile range
  - Large SD in relation to the mean
    - **(Mean – Minimum)/SD**: < 2 indicates skewed data
    - **(Maximum – Mean)/SD:** < 2 indicates skewed data

- **Addressing skew**
  - Get statistical advice before proceeding
    - May be no action required
    - Possible actions may include sensitivity analysis without skewed studies, log transformation or other methods

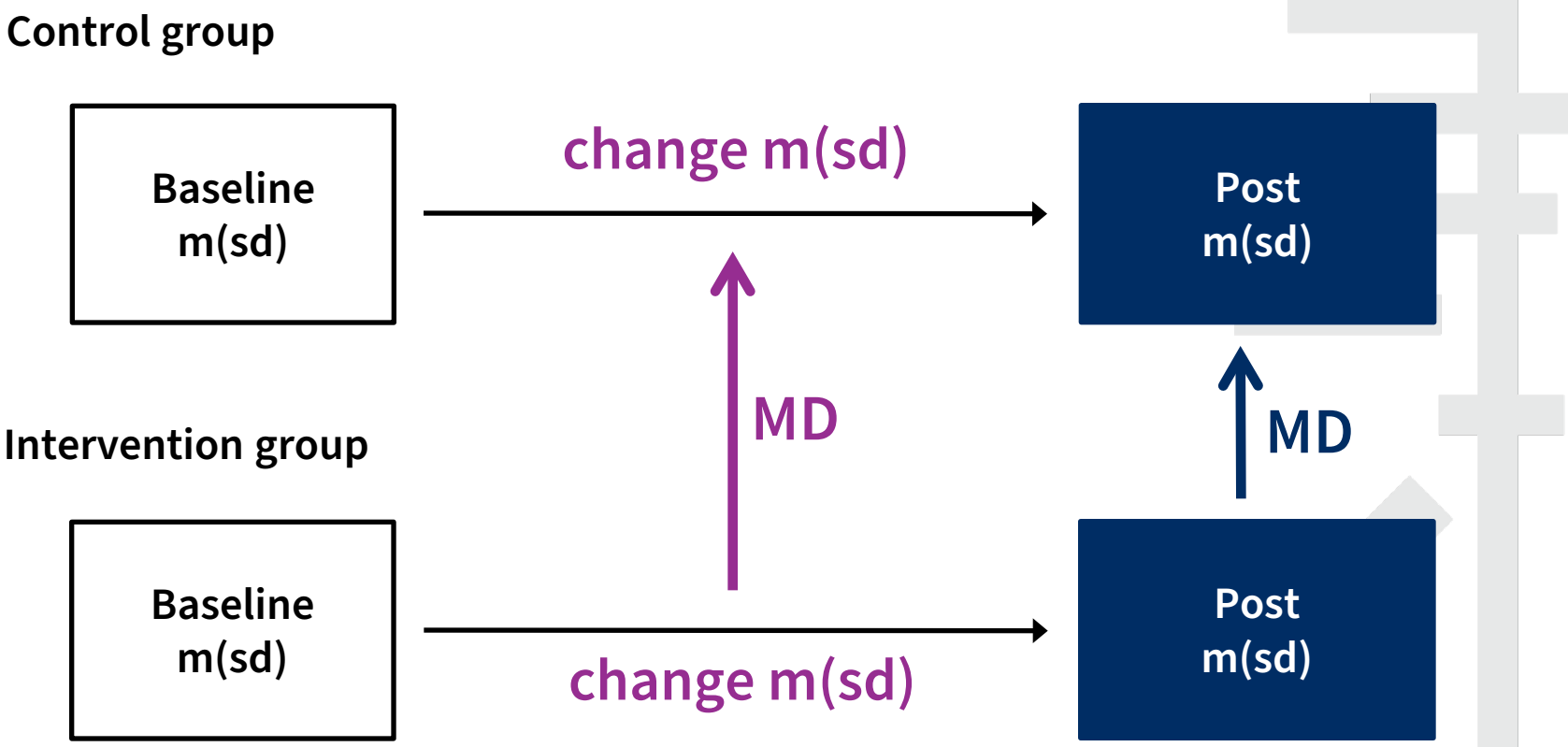
Collecting continuous outcome data

- Six numbers needed for meta-analysis

	Mean	SD	N
Intervention	20	9.1	65
Control	33	8.6	67

Try to collect the actual number measured for each outcome, at each time point

Post-intervention vs change from baseline



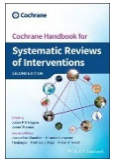
Post-intervention vs change from baseline

- Can you combine studies using both in one meta-analysis?
  - **For MD:** Yes, since both estimate the same intervention effect
  - **For SMD:** No, since SDs differ systematically between post-intervention and change from baseline scores
- ANCOVA (adjusting for baseline) is statistically preferred if reported
- Work with what is reported in your included studies:
  - Either post-intervention or change scores can be used
  - Can use a mixture (only for MD, not SMD)
  - Better to be consistent if possible
  - Avoid selective outcome reporting
  - Change scores require SD of the change

Other data formats can also be used

- **Statistics other than mean and SD**
  - e.g. standard error, confidence interval, P value, t value, median, interquartile range
  - Clarify with the author if unclear
  - Can often calculate or estimate the mean and SD

- **Overall effect estimate**
  - e.g. MD, ANCOVA, ratio of means, ratio of geometric means
  - Can include in meta-analysis using generic inverse variance method
  - Need a measure of variance (e.g. SE, 95% CI)



See Section 6.5.2 of the Handbook

What to include in your protocol

- Specify effect measures to be used for dichotomous and continuous outcome data
- **For continuous outcome data, specify:**
  - Plans for handling different scales
  - Whether you have a preference to use post-intervention or change scores
  - Plans for converting statistics to the required formats
  - Minimally Important Differences if possible

Cochrane Training

Take home message: dichotomous

- Risks and odds are two ways of expressing how likely an event is
- Risk ratio, odds ratio and risk difference compare chance between two groups
- To enter dichotomous data, you need the **number of events** and the **total number** in each group

Take home message: continuous

- Mean difference and standardised mean difference compare continuous measures between two groups
- For basic analysis of continuous data, you need the **mean** and **SD** and the **number of participants** in each group
- Both change and post-intervention data can be combined in your analysis
- The required statistics can often be calculated from the reported data

References

- Higgins JPT, Li T, Deeks JJ (editors). **Chapter 6: Choosing effect measures and computing estimates of effect.** In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). Cochrane, 2020. Available from www.training.cochrane.org/handbook.
- 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 6.1 (updated September 2020). Cochrane, 2020. Available from www.training.cochrane.org/handbook.

Acknowledgements

- Compiled by Miranda Cumpston
- Based on materials by Sally Hopewell, Phil Alderson, Jon Deeks, Deborah Caldwell, the Cochrane Statistical Methods Group, Cochrane Australia and Cochrane UK
- Approved by the Convenors of Cochrane Methods Groups