UNIVERSITÄT BERN

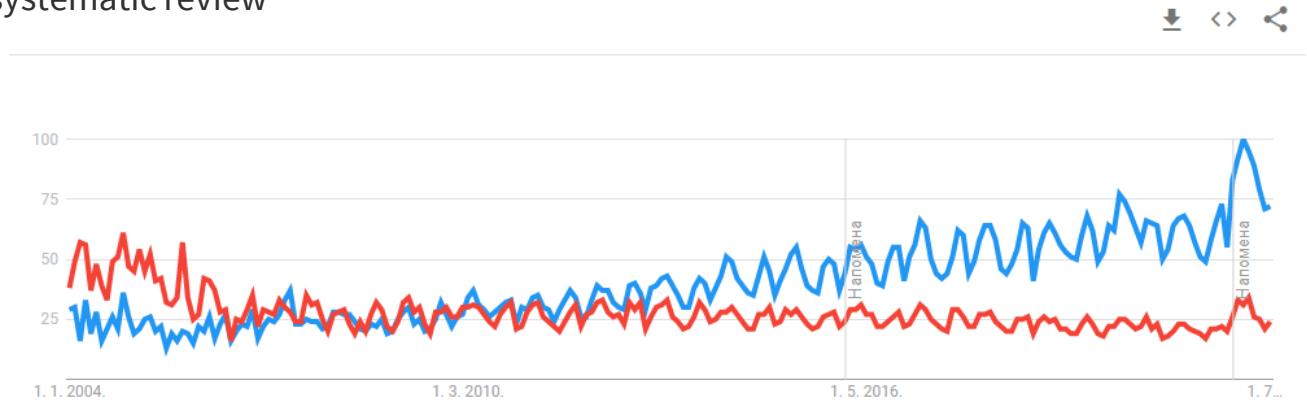
Why do we need systematic reviews and meta-analyses?

PD Marija Glisic, MD, M.H.S., PhD

Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland & Swiss Paraplegic Research, Nottwil

Google trends 2014-2022

Red: meta-analysis
Blue: systematic review



Value of 100 represents the highest popularity for given term

Purposes of Systematic Review

- To provide a clear and comprehensive overview of available evidence on a given topic.
- To identify research gaps in our current understanding of a field
- To highlight methodological concerns in research studies that can be used to improve future work in the topic area
- To identify questions for which the available evidence provide clear answers and thus for which further research is not necessary

Purposes of Meta-analysis

- Increase power and improve precision of estimates of effect from numerous smaller studies by pooling the total sample sizes across a number of investigations
- Explain contradictory results (heterogeneity) that arise between studies conducted in different settings
- Conduct subgroup analyses that could not be conducted or were not thought of in the initial studies and help generate new hypotheses

Real life example

Endocrinological and inflammatory markers in individuals with spinal cord injury: A systematic review and meta-analysis

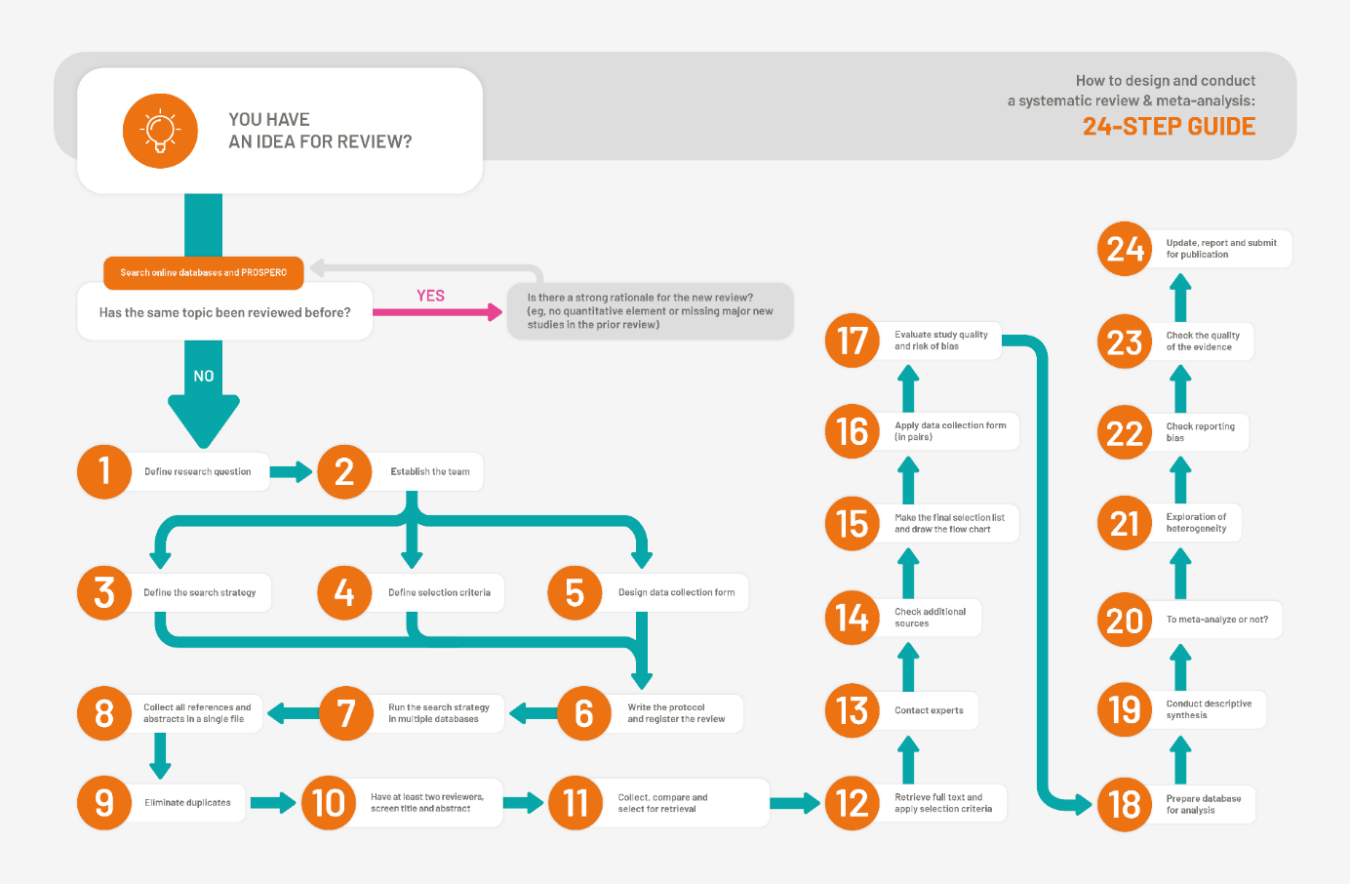
Springer

Table 2 Weighted mean difference of biomarkers among spinal cord injury and able-bodied population

Outcome (units)	Studies which reported higher levels in SCI	Studies which reported lower levels in SCI	No association	Number of studies	SCI, N	ABI, N	Weighted Mean Difference	95% confidence interval	I^2 test for heterogeneity	x^2 test for heterogeneity ($p^{\chi 2}$)
Inflammatory markers										
CRP (mg/L)	[11, 23, 37, 84–86]	-	[87, 88]	8	330	249	2.79	**1.75, 3.83***	87.3%	**<0.001***
hsCRP (mg/L)	[89]	-	[90]	2	114	113	0.07	-0.07, 0.20	66.3%	0.085
IL-6 (pg/mL)	[11, 91, 92]	-	[24]	4	150	104	2.52	**1.82, 3.21***	81.1%	**0.001***
TNF alpha (pg/mL)	[91]	-	[24]	2	68	45	18.51	-26.14, 63.15	85.0%	**0.010***
Insulin										
Insulin (pmol/L)	[93, 94]	[95]	[11, 87, 89, 90, 96–101]	13	493	485	3.99	-2.84, 10.83	50.5%	**0.019***
Glucose (mmol/L)	[102]	[93, 96, 103]	[11, 86–90, 94, 95, 97–101, 104–108]	22	1073	1003	-0.08	-0.19, 0.03	83.8%	**<0.001***
Creatinine										
Creatinine(μmol/L)	-	[11, 109–111]	[57, 84, 108]	7	260	136	-14.23	**-21.57, -6.89***	90.3%	**<0.001***
Vitamin D										
25(OH)D (nmol/L)	-	[60, 84]	[56, 57, 111]	5	289	123	-10.32	**-20.47, -0.18***	57.2%	0.053
1,25(OH)D (pmol/L)	[56]	-	[57]	2	140	64	6.82	-50.10, 63.73	96.8%	**<0.001***
Hormone and Growth Factors										
Total Testosterone (nmol/L)	-	[22, 47, 111–117]	[26, 65, 66, 110, 119–122]	18	601	512	-2.61	**-4.42, -0.79***	89.6%	**<0.001***
Free Testosterone (nmol/L)	-	[25, 115, 117]	[26, 66, 123]	6	141	123	-0.01	-0.024, 0.004	81.8%	**<0.001***
TSH (mU/L)	-	-	[62, 65, 124]	3	135	99	-0.03	-0.26, 0.20	0.0%	0.490
T3 (nmol/L)	-	[125]	[65, 124, 126]	4	108	100	-0.05	-0.29, 0.19	82.4%	**0.001***
LH (IU/L)	[26]	[22, 47]	[65, 66, 110, 111, 114, 116, 117, 119–123, 126]	15	452	365	0.27	-0.61, 1.14	85.6%	**<0.001***
FSH (IU/L)	[65, 110, 111, 121]	[22, 47]	[66, 114–117, 119, 123]	13	388	305	0.59	-1.24, 2.43	95.7%	**<0.001***
GH (μg/L)	-	-	[65, 66, 98]	3	48	42	-0.32	-0.67, 0.04	0.0%	0.550
Cortisol (nmol/L)	[43, 127]	-	[65, 126]	4	90	76	103.43	**10.75, 196.11***	67.5%	**0.026***
ACTH (pmol/L)	-	-	[126, 127]	2	36	28	1.59	-0.40, 3.58	0.0%	0.786
Adiponectin (μg/mL)	-	-	[87, 89]	2	29	29	0.92	-3.12, 4.96	24.6%	0.249
Aldosterone (pmol/l)	-	-	[128, 129]	2	32	22	100.62	-69.57, 270.81	40.2%	0.196
IGF-1 (nmol/L)	-	[65]	[66, 84, 97]	4	76	59	-6.82	**-9.24, -4.40***	0.0%	0.529
Leptin (nmol/L)	[67, 89, 94, 97, 130, 131]	-	[98]	7	176	141	0.19	**0.10, 0.27***	53.1%	**0.047***
Prolactin (μg/L)	-	-	[22, 47, 65, 121, 126, 127]	6	209	134	0.80	-1.01, 2.61	76.5%	**0.001***

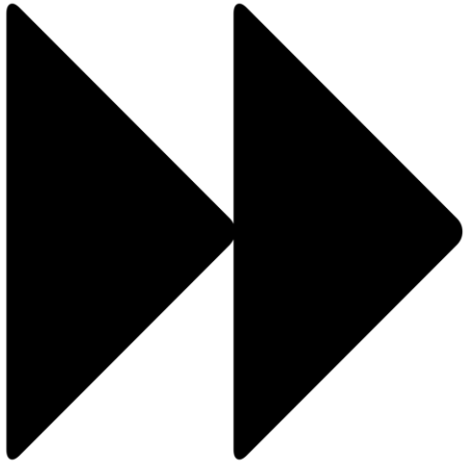
Reviews in Endocrine and Metabolic Disorders

Detailed steps to systematic reviews:
“The 24”



Fast forward mode towards understanding meta-analysis

5 day-course → 3-hour-course



First:

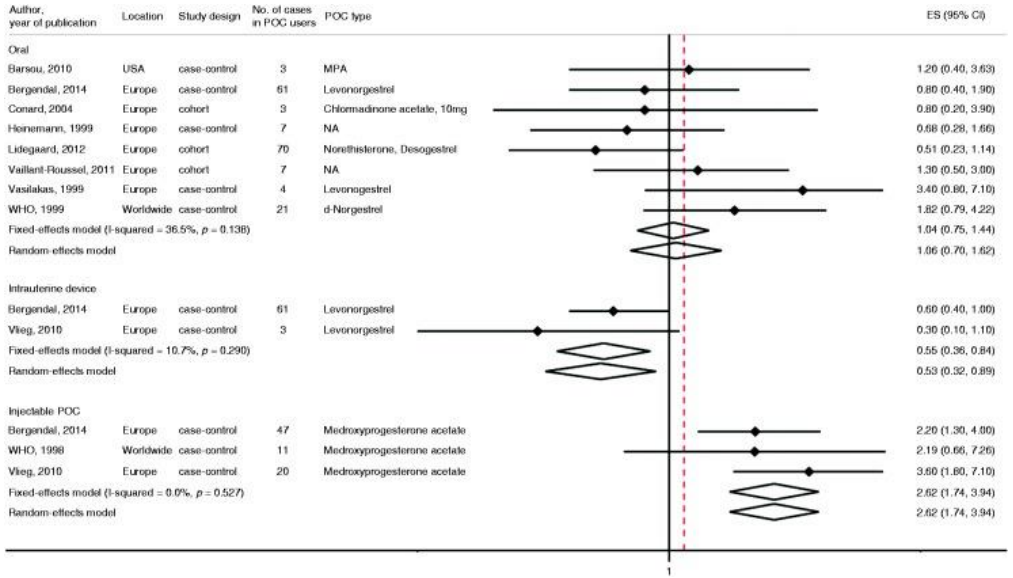
Decide whether narrative/descriptive data synthesis or meta-analysis is suitable

- How many studies are sufficient for meta-analysis?
- What information is essential for meta-analysis?
- When is heterogeneity too large?
- Do "study quality" and methodological rigor determine whether to meta-analyze the evidence?
- Does study design determine whether to meta-analyze the evidence?

How many studies are sufficient for meta-analysis

- Meta-analysis is possible if association estimates from two studies are available

The association between progestin-only contraceptive (POC) use and risk of venous thromboembolism by route of administration



What information is essential for meta-analysis?

- To combine study results, measurements of association estimates, such as relative risk or odds ratio), from individual studies and standard errors or 95% confidence intervals (CIs) of the estimate are needed

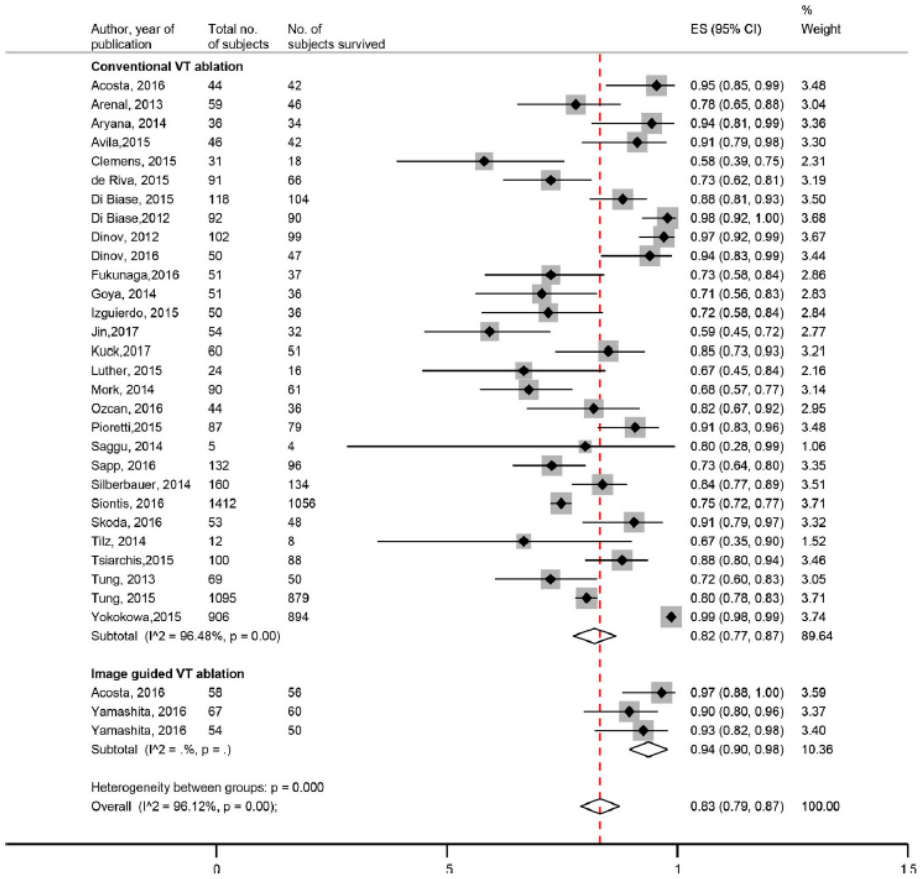


Fig. 3 Forest plot survival—image-guided versus non-image-guided. (*CI* confidence interval, *ES* effect size, *VT* ventricular tachycardia)

When is heterogeneity too large?

- Subjective
- No widely accepted, automated quantitative measures to grade it
- Deciding to perform meta-analyses should not be based solely on statistical heterogeneity

To decide on non-statistical heterogeneity (e.g., clinical heterogeneity), authors should review the included studies and check whether:

1. Studies used different methods to define exposure and/or outcome
2. Different study designs were used
3. Different analyses and methods were applied to generate the estimates
4. Different estimates with different levels of confounding adjustments were reported across studies
5. There were variations in populations included across different studies and
6. Studies differ by their quality/risk of bias
7. Few studies are available to make a meaningful analysis that could account for the differences across studies

Do "study quality" and methodological rigor determine whether to meta-analyze the evidence?

- "Study quality" is a complex term
- Involves assessing methodological rigor (what was done) and completeness or accuracy of reporting (what is reported to have been done) within individual studies.

Table 2. The subgroup analyses by study characteristics

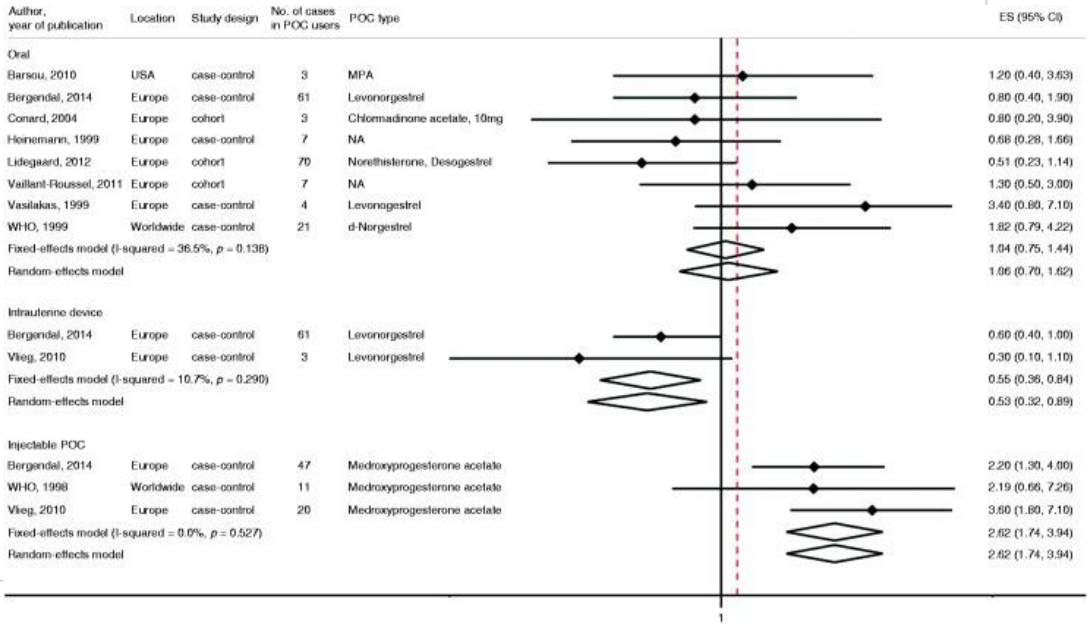
Subgroups by Study Characteristics		Number of studies	[1]Difference, Mean (95 % CI)	[2]I² for heterogeneity	[3]P value for heterogeneity
Phytoestrogen use and mean body weight change					
[a]Primary study goal of the RCT	Body composition	7	-0.04 (-0.68; 0.60)	68.7%	0.59
	Other	11	-0.21 (-0.64; 0.23)	62.3%	
[b]Median years since menopause onset	≤4.1 years	4	-0.23 (-0.6; 0.14)	16.5%	0.87
	>4.1 years	8	0.07 (-0.51; 0.64)	68.2%	
	Unknown	6	-0.29 (-0.63; 0.04)	0%	
[c]Route of administration	Tablet/capsule	10	-0.17 (-0.82; 0.47)	76.4%	0.52
	Diet	8	-0.17 (-0.42; 0.07)	0%	
[d]Intervention type	Soy products	6	-0.49 (-1.21;0.23)	65.4%	0.83
	Isoflavone mix	10	-0.24(-0.46; -0.01)	0%	
	Daidzein	2	0.92 (0.24; 1.59)	55%	
[e]Median number of study participants	≤66women	9	-0.49 (-0.87; -0.11)	0%	0.22
	>66 women	9	0.11 (-0.31; 0.52)	75.2%	
[f]Intervention duration	≤24 weeks	15	-0.10 (-0.51; 0.30)	69.5%	0.9
	> 24weeks	3	-0.32 (-1.02; 0.37)	53.2%	
[g]Location	Asia	7	0.22 (-0.25; 0.69)	76.8%	0.48
	Europe	4	-0.35 (-0.78; 0.07)	11%	
	North America	4	-0.6 (-1.44; 0.24)	0%	
	South America	3	-0.48 (-1.13; 0.17)	0%	
[h]Risk of bias	High	2	-0.85 (-1.93; 0.22)	0%	0.39
	Low to medium	16	-0.09 (-0.45; 0.27)	69.6%	

MANUSCRIPT

Does study design determine whether to meta-analyze the evidence?

- Including all study designs in systematic reviews reduces subjective interpretations of potential biases and inappropriate study exclusions
- The decision to meta-analyze results across all study designs depends on research questions

The association between progestin-only contraceptive (POC) use and risk of venous thromboembolism by route of administration



Descriptive data synthesis

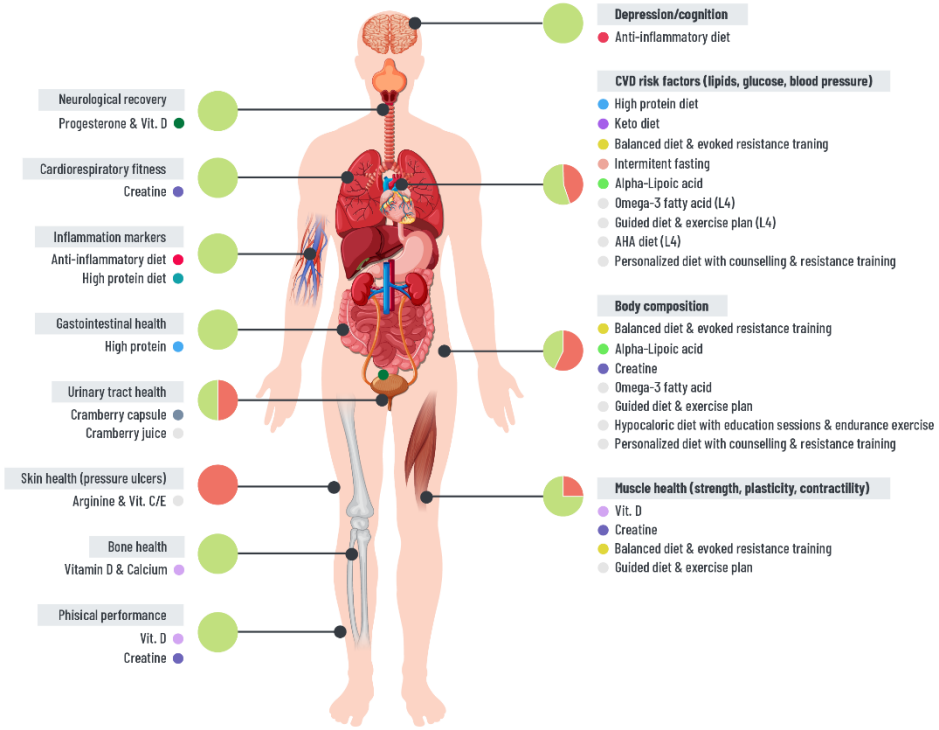


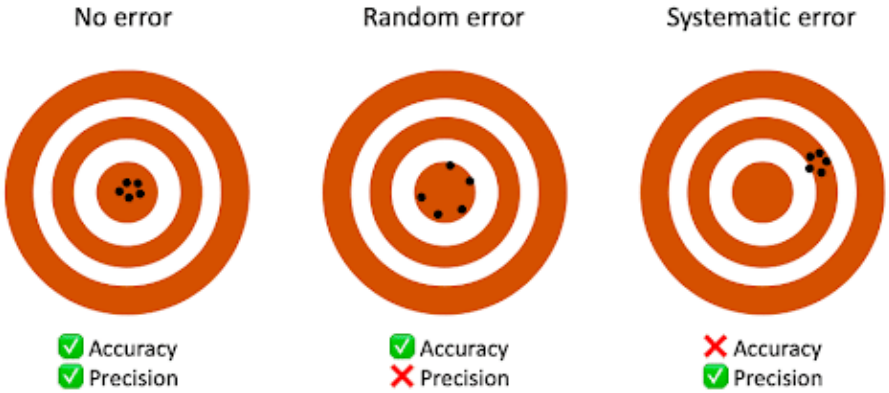
Table 1. Steps for synthesizing data effectively

Step 1	Group studies	*Choose an appropriate grouping rationale.* A. PECO [population (male only participants vs mixed population; healthy vs individuals with comorbidities; animal vs human evidence)], exposure/comparison and outcome (reported on continuous vs dichotomized scale) B. Study design (cross-sectional vs longitudinal studies) C. Risk of bias (low quality vs moderate or high-quality evidence) D. Association estimates: consider type (beta coefficients, risk ratios, odds ratios, hazard ratios, etc.) and direction of association (higher risk in exposed population vs no association). To accurately interpret *p-values* and 95% confidence intervals, identify and understand the direction of associations.
Step 2	Follow the same synthesis consistently	A. Create additional tables using study groupings to find patterns among studies. Provide separate tables for cross-sectional and observational studies, for example. B. Convert association estimates if possible. For comparison among studies, convert odds ratios to standardized mean differences (72). C. Present most interesting findings using graphical methods, such as arrows indicating increased or decreased risk between groups. D. If meta-analysis is not possible, use the data extraction sheet to conduct minimal statistical analyses. For example, calculate total numbers of study participants, mean age, mean number of male participants, or other relevant study population, exposure, or outcome characteristic.
Step 3	Report findings clearly	A. Use appropriate language. B. Keep reporting style uniform across results section. For instance, if studies are grouped, start with a paragraph explaining grouping variables. C3. Provide summary tables and/or figures to support findings reported in results section.
Step 4	Discuss findings objectively	*Summarizing what best reflects reviewed evidence can be challenging.* D1. Report based on grouping parameters from Step 1. Graphical summaries support interpreting findings (especially when analyzing many methodologically different studies). D2. Discuss methodological strengths and limitations of reviewed evidence. For example, levels of adjustment across studies, heterogeneity that precluded quantitative synthesis, or risks of bias. D3. Identify literature gaps and provide directions for future research

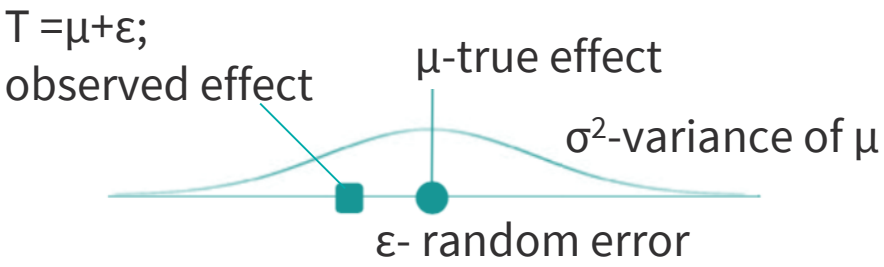
Second:
Understand the concept of meta-analysis and different models

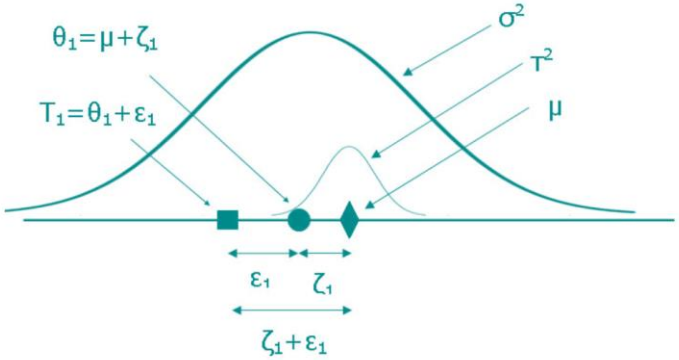
- Meta-analysis mathematically combines different study results, it computes summary statistics for each study, then further summarizes and interprets study-level summary statistics

Fixed vs. random effect



Fixed vs. random effect





Fixed vs. random effect

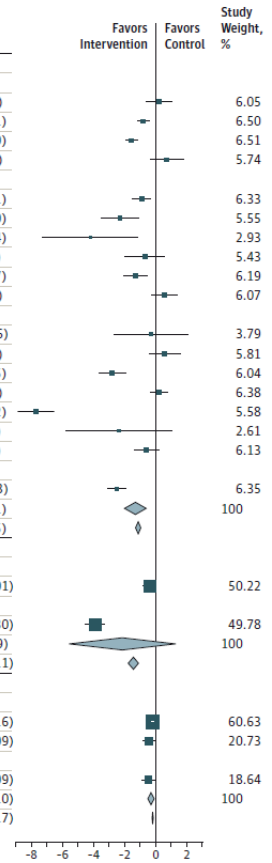
	Fixed effect model	Random effects model
Assumption	All studies share a common true overall effect. The differences in observed effect between studies is due to random noise.	Each study has an effect of its own, where each study effect being the combination of a true overall effect, plus a difference coming from between-study variation and random noise.
True effect	One true effect.	True effect varies.
Interpretation of pooled estimate	Pooled estimate is common effect.	Pooled effect is mean effect of all true study effects.
Goal of study	Goal is to estimate the true effect for the identified population.	Goal is to estimate the mean of all true effect sizes.
Variability accounted for	Accounts for within study variability. Between-study variability is ignored.	Accounts for within study and between study variability.
When to pool the effects?	All study characteristics are identical to a large degree.	Studies have enough in common to statistically pool the information, but it is not required that these studies have an identical true effect size.
Interpretation of target population	Inference is relevant for the (homogenous) population included in the study.	Inference is relevant for the (heterogenous) population types included in the study.
Implication	Large studies are more influential (receive more weight) in analysis as compared to random effects model.	Large studies are less influential (receive less weight) in analysis as compared to fixed effect model: a more equal distribution of weights.
Under heterogeneity	More likely to lead to incorrect statistically significant findings than in the random effects model.	Larger p-values and wider confidence intervals than in the fixed effect model.
Without heterogeneity	Correct statistical inference.	Correct mean estimate, with too wide confidence intervals and too large a p-value.
Models for continuous outcomes	Inverse-variance fixed effect	Inverse-variance random effects (DerSimonian-Laird)
Models for binary outcomes		
○ Risk ratio	Inverse-variance fixed effect; Mantel-Haenszel	Inverse-variance random effects (DerSimonian-Laird)
○ Odds ratio	Inverse-variance fixed effect; Mantel-Haenszel; Peto	Inverse-variance random effects (DerSimonian-Laird)
○ Risk difference	Inverse-variance; Mantel-Haenszel	Inverse-variance random effects (DerSimonian-Laird)

Fixed vs. random effect

Example: the effect of phytoestrogens on menopausal symptoms

Figure 2. Meta-analysis of Randomized Clinical Trials on the Associations Between Use of Phytoestrogen Supplementation and Menopausal Symptoms

Source	No. of Participants		Change, Mean (95% CI)[a]		Difference, Mean (95% CI)[b]	Study Weight, %
	Intervention	Control	Intervention	Control		
No. of Hot Flashes in 24 Hours						
Dietary soy isoflavones						
Lewis et al,[47] 2006	33	33	-0.71 (-3.79 to 2.37)	-0.93 (-4.67 to 2.81)	0.22 (-0.62 to 1.06)	6.05
Cheng et al,[33] 2007	30	30	-0.80 (-2.47 to 0.87)	0.0 (-1.37 to 1.37)	-0.80 (-1.19 to -0.41)	6.50
Albertazzi et al,[26] 1998	51	53	-5.01 NR	-3.42 NR	-1.59 (-1.95 to -1.20)	6.51
Van Patten et al,[59] 2002	78	79	-1.8 (-7.01 to 3.41)	-2.5 (-10.4 to 5.40)	0.70 (-0.37 to 1.77)	5.74
Supplements and extracts of soy isoflavones						
Aso et al,[28] 2012	77	83	-1.9 (-5.43 to 1.63)	-1 (-4.92 to 2.92)	-0.90 (-1.49 to -0.31)	6.33
Nahas et al,[51] 2007	40	40	-6.5 (-11.4 to -1.64)	-4.2 (-10.0 to 1.60)	-2.30 (-3.50 to -1.10)	5.55
Faure et al,[39] 2002	39	36	-6.4 (-18.6 to 5.83)	-2.2 (-16.3 to 11.9)	-4.20 (-7.26 to -1.14)	2.93
Penotti et al,[52] 2003	28	34	-5.3 (-10.6 to 0.01)	-4.6 (-9.30 to 0.10)	-0.70 (-1.98 to 0.58)	5.43
Ferrari et al,[40] 2009	85	95	-3.7 (-8.8 to 1.40)	-2.4 (-7.01 to 2.21)	-1.30 (-2.03 to -0.57)	6.19
Hachul et al,[41] 2011	19	19	-4.33 (-6.92 to -1.74)	-4.9 (-7.41 to -2.39)	0.57 (-0.26 to 1.40)	6.07
Red clover						
Knight et al,[44] 1999	13	12	-3.1 (-9.27 to 3.07)	-2.8 (-8.44 to 2.84)	-0.30 (-2.66 to 2.06)	3.79
Baber et al,[30] 1999	25	26	-1.18 (-4.98 to 2.62)	-1.77 (-5.26 to 1.72)	0.59 (-0.43 to 1.61)	5.81
Jeri et al,[43] 2002	30	30	-3.4 (-6.97 to 0.17)	-0.60 (-3.58 to 2.38)	-2.80 (-3.65 to -1.95)	6.04
Atkinson et al,[29] 2004	102	103	-0.8 (-4.92 to 3.32)	-1.0 (-4.53 to 2.53)	0.20 (-0.34 to 0.74)	6.38
Lipovac et al,[24] 2012	50	59	-8.6 (-14.3 to -2.92)	-0.9 (-7.51 to 5.71)	-7.70 (-8.88 to -6.52)	5.58
van de Weijer et al,[58] 2002	16	14	-2.08 (-8.14 to 3.98)	0.29 (-11.0 to 11.6)	-2.37 (-5.75 to 1.01)	2.61
Tice et al,[56] 2003	84	85	-3.4 (-9.20 to 2.40)	-2.8 (-7.05 to 1.45)	-0.60 (-1.38 to 0.18)	6.13
Other phytoestrogens						
Dánna et al,[36] 2007	198	191	-2.5 (-8.09 to 3.09)	0.0 (-5.68 to 5.68)	2.50 (-3.07 to -1.93)	6.35
Random effects					-1.31 (-2.02 to -0.61)	100
Fixed effects					-1.12 (-1.29 to -0.95)	
No. of Night Sweats in 24 Hours						
Dietary soy isoflavones						
Cheng et al,[33] 2007	30	30	-0.60 (-2.21 to 1.01)	-0.20 (-1.63 to 1.23)	-0.40 (-0.79 to -0.01)	50.22
Red clover						
Lipovac et al,[24] 2012	50	59	-3.9 (-6.84 to -0.96)	0.0 (-3.37 to 3.37)	-3.90 (-4.50 to -3.30)	49.78
Random effects					-2.14 (-5.57 to 1.29)	100
Fixed effects					-1.44 (-1.77 to -1.11)	
Vaginal Dryness Score						
Supplements and extracts of soy isoflavones						
Kotsopoulos et al,[46] 2000	44	50	-0.35 (-0.59 to -0.11)	-0.15 (-0.35 to 0.05)	-0.20 (-0.24 to -0.16)	60.63
Colacurci et al,[34] 2004	15	15	-0.54 (-1.48 to 0.40)	-0.08 (-1.16 to 1.00)	-0.46 (-0.83 to -0.09)	20.73
Other phytoestrogens						
Brzezinski et al,[31] 1997	78	36	-0.93 (-3.03 to 1.17)	-0.44 (-2.34 to 1.46)	-0.49 (-0.89 to -0.09)	18.64
Random effects					-0.31 (-0.51 to -0.10)	100
Fixed effects					-0.21 (-0.25 to -0.17)	



Third:
Follow the statistical analysis plan

When preparing statistical analysis plans, consider how the plan answers the following questions:

1. What is the direction and size of the summary association estimate?
2. What is the statistical heterogeneity?
3. Does a single study drive the overall association estimate?
4. What study and study population characteristics influence the size of pooled association estimate?
5. Are the results reliable?
6. Is there a serious risk from reporting bias, such as publication or selective?

Fourth:
Prepare datasets for meta-analysis

- Convert units for consistency before combining and meta-analyzing.
- Resolve differences in reporting summary statistics, such as measures of central tendency and spread
- Pool multiple intervention groups
- Map adjusted variables

UNIVERSITÄT BERN

Meta-analysis of continuous and binary data

PD Marija Glisic, MD, M.H.S., PhD

Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland & Swiss Paraplegic Research, Nottwil

24-STEP GUIDE

Data needed for meta-analysis

Box 1. Information needed to meta-analyze specific association estimates

Association estimates	Extracted information from original studies
Prevalence	Study name (ID) Study size Number of cases Prevalence 95% confidence interval
Odds ratio **Hazards ratio** **Relative risk**	Study name (ID) Number of cases Number of events in cases Number of non-events in cases Number of controls Number of events in controls Number of non-events in controls
Correlation coefficient	Study name (ID) Study size Correlation coefficient 95% confidence interval
Beta from regression coefficient	Study name (ID) Study size Beta coefficient 95% confidence interval
Mean difference	Study name (ID) Number of individuals in group 1 Mean value in group 1 Standard deviation of the mean in group 1 Number of individuals in group 2 Mean value in group 2 Standard deviation of the mean in group 2

Continuous Data

- Weighted mean difference (WMD)
- Standardized mean difference (SMD)

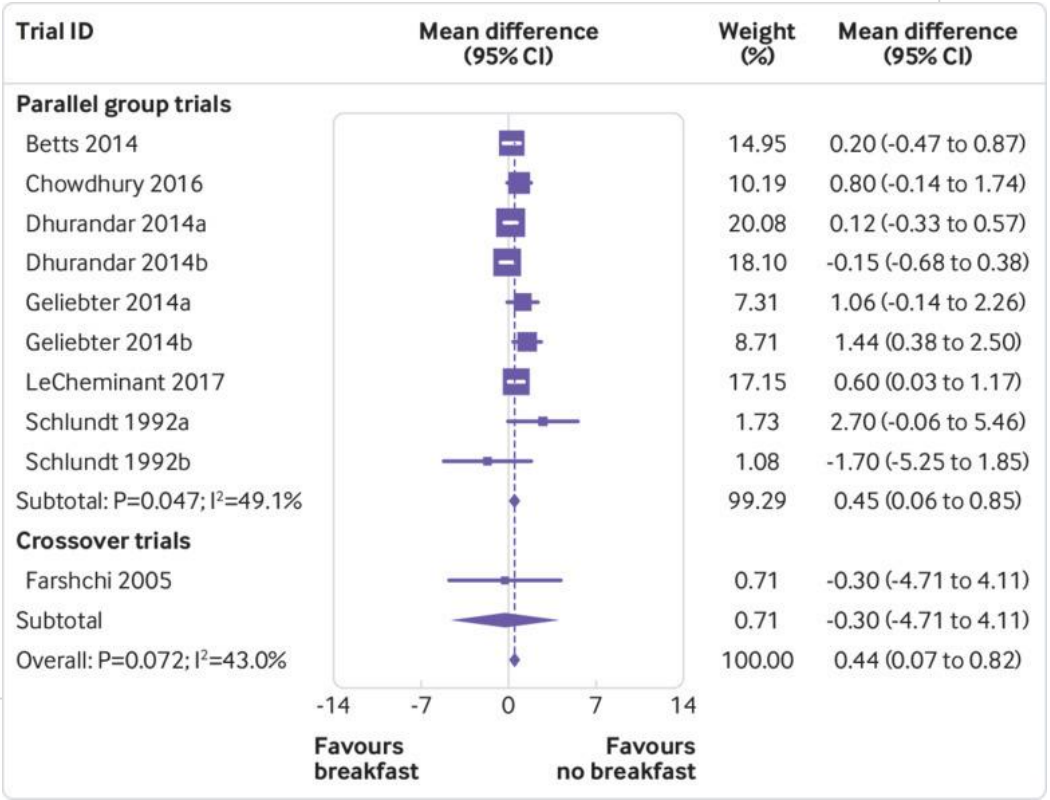
Table S4. Meta-analysis for continuous outcomes using Stata (metan) and R (meta/metacont)

	Stata	R
Data input (for each study)	Size of group 1 and 2: Mean effect size in group 1 and 2: Standard deviation of effect in group 1 and 2:	n1, n2 m1, m2 sd1, sd2
Method	Inverse variance weighting	
Assumptions	Normality in effect sizes within trial arms. Between trial variations in standard deviations are attributed to differences in precision, and are assumed equal in both study arms.	
Fixed effect estimate	`metan n1 mean1 sd1 n2 mean2 sd2, nostandard fixed`	`metacont (n1, m1, sd1, n2, m2, sd2, comb.fixed=TRUE, comb.random=FALSE, sm="MD", data=dat)`
Random effects estimate	`metan n1 mean1 sd1 n2 mean2 sd2, nostandard random`	`Metacont (n1, m1, sd1, n2, m2, sd2, comb.fixed=FALSE, comb.random=TRUE, sm="MD", data=dat)`
Random effects estimate, standardized mean differences (Hedges g)	`metan n1 mean1 sd1 n2 mean2 sd2, hedges random`	`metacont(n1, m1, sd1, n2, m2, sd2, comb.fixed=FALSE, comb.random=TRUE, sm="SMD", method.smd="Hedges", data=dat)`

Continuous Data

- Weighted mean difference (WMD)
- Standardized mean difference (SMD)

Random effects meta-analysis of the mean difference in weight (kg), based on breakfast consumption or no breakfast consumption.

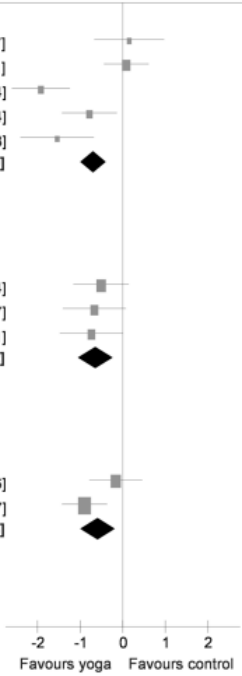


Continuous Data

- Weighted mean difference (WMD)
- Standardized mean difference (SMD)

Depression

Comparison / Study	Yoga Mean	Yoga SD	Yoga Total	Control Mean	Control SD	Control Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Yoga vs. Usual Care								
Butler et al. (2008)	14.25	8.99	12	12.9	8.14	11	13.5%	0.15 [-0.67, 0.97]
Field et al. (2012b)	20.12	10.51	28	19.27	10.12	28	33.0%	0.08 [-0.44, 0.61]
Khumar et al. (1993)	57.52	13.49	25	77.92	6.04	25	19.6%	-1.92 [-2.60, -1.24]
Sahidi et al. (2011)	10	6.9	20	15.2	6.1	20	21.7%	-0.78 [-1.43, -0.14]
Woolery et al. (2004)	3.9	4.66	13	11	4.32	15	12.2%	-1.54 [-2.40, -0.68]
Subtotal (95% CI)			**98**			**99**	**100.0%**	**-0.69 [-0.99, -0.39]**
Heterogeneity: Chi² = 28.81, df = 4 (P < 0.00001); I² = 86%								
Test for overall effect: Z = 4.48 (P < 0.00001)								
Yoga vs. Relaxation								
Lavretsky et al. (2012)	-7.4	3.7	23	-5.3	4.5	16	39.4%	-0.51 [-1.16, 0.14]
Rohini et al. (2000)	11.1	5.2	15	15.6	7.7	15	30.5%	-0.67 [-1.40, 0.07]
Sharma et al. (2005)	8.27	4.37	15	11.53	4.26	15	30.1%	-0.74 [-1.48, 0.01]
Subtotal (95% CI)			**53**			**46**	**100.0%**	**-0.62 [-1.03, -0.22]**
Heterogeneity: Chi² = 0.22, df = 2 (P = 0.90); I² = 0%								
Test for overall effect: Z = 3.01 (P = 0.003)								
Yoga vs. Aerobic Exercise								
Sahidi et al. (2011)	10	6.9	20	11.1	6.2	20	42.3%	-0.16 [-0.79, 0.46]
Veale et al. (1992)	13.31	1.86	22	14.96	1.8	46	57.7%	-0.90 [-1.43, -0.37]
Subtotal (95% CI)			**42**			**66**	**100.0%**	**-0.59 [-0.99, -0.18]**
Heterogeneity: Chi² = 3.08, df = 1 (P = 0.08); I² = 68%								
Test for overall effect: Z = 2.85 (P = 0.004)								
Test for subgroup differences: Chi² = 0.17, df = 2 (P = 0.92), I² = 0%								



Binary Data

- Proportion and prevalence
- Risk ratio
- Odds ratio
- Relative risk
- Hazard ratio
- Risk differences

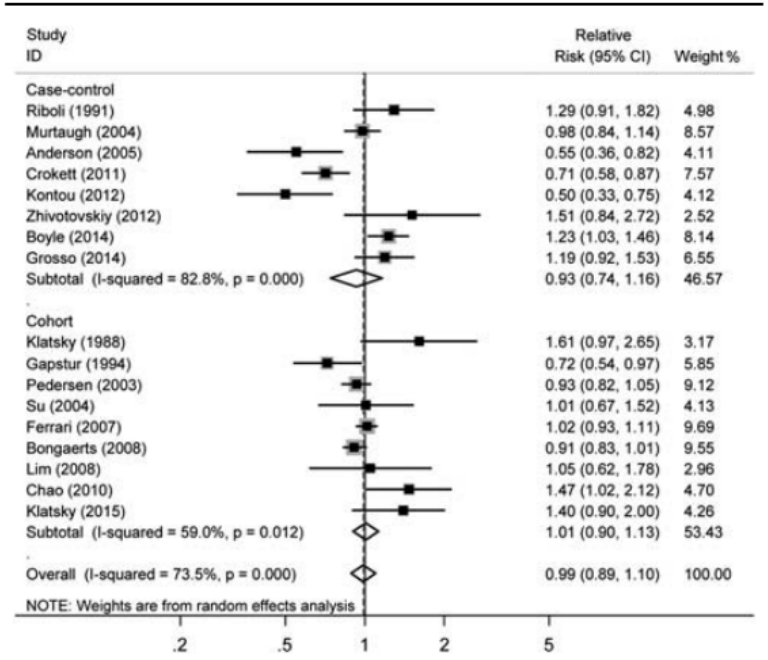
Table S3. Overview on effect sized measures for dichotomous data, with normalizing transformation and example codes in Stata (metan) and R (meta/metacont)

	Risk ratio	Odds ratio	Risk difference
Effect size Where: P_{treat}, P_{contr}: Probability of event in treatment and control arm n1, n2: number of events in treatment and control arm N1, N2: sample size in treatment and control arm For the odds ratio notation we use: n1=A; N1-n1=B; n2=C; N2-n2=D	Ratio of the probability of an event of both arms: $ES=P_{treat}/P_{contr}$ =(n1/N1) / (n2/N2)	Ratio of the odds of the event in one arm to the odds of the event in the other arm: $ES= Odds_{treat}/Odds_{contr}$ = A*D / B*C	Difference in risks of an event of both arms: $ES=P_{treat} – P_{contr}$ =(n1/N1) – (n2/N2)
Interpretation	Risk ratio of 1 means that the risk is the same in both arms, indicating no effect, while a risk ratio less or greater than 1 would mean that the risk was relatively lower or higher in one or the other arm, respectively	Odds ratio of 1 means that the odds is the same in both arms, indicating no effect, while a odds ratio less or greater than 1 would mean that the odds was relatively lower or higher in one or the other arm, respectively	Risk difference of 0 means that the risk is the same in both arms, indicating no effect, while a negative of positive risk difference would mean that the risk was relatively lower or higher in one or the other arm, respectively
Usage	Used in prospective studies (cohort studies). Cannot be used in retrospective (case control) studies	Commonly used as output of logistic regression analyses	Less common than risk ratio or odds ratio, but better understandable
Transformation	`TES=log(ES)`	`TES=log(ES)`	`TES=ES` `(no transformation needed)`

Binary Data

- Proportion and prevalence
- Risk ratio
- Odds ratio
- Relative risk
- Hazard ratio
- Risk differences

Fig. 2



Meta-analysis of wine consumption and the risk of colorectal cancer for any drinkers versus nondrinkers. CI, confidence interval.

Binary Data

- **Fixed-effect methods** (Mantel-Haenszel, Peto and inverse variance)
- **Random-effects method** (DerSimonian and Laird inverse variance).

Binary Data

- Double- and single-zero studies
- Continuity correction
- Generalized linear mixed models
- Beta-binomial regression

metan package in Stata
and the metabin command from the meta
library in R correct by these by default

Can we combine continuous and binary data?

Chinn et al. 2000

$$\mathrm{SMD} = \frac{\sqrt{3}}{\pi} \ln \mathrm{OR}.$$

- SE of the log odds ratio can be converted to the standard error of a SMD by multiplying by the constant of 0.5513
- SMDs can be re-expressed as log odds ratios by multiplying by $\pi/\sqrt{3}=1.814$.
- Once SMDs (or log odds ratios) and their standard errors have been computed for all studies in the meta-analysis, they can be combined using the generic inverse-variance method.

A Hands-On Guide by Peter