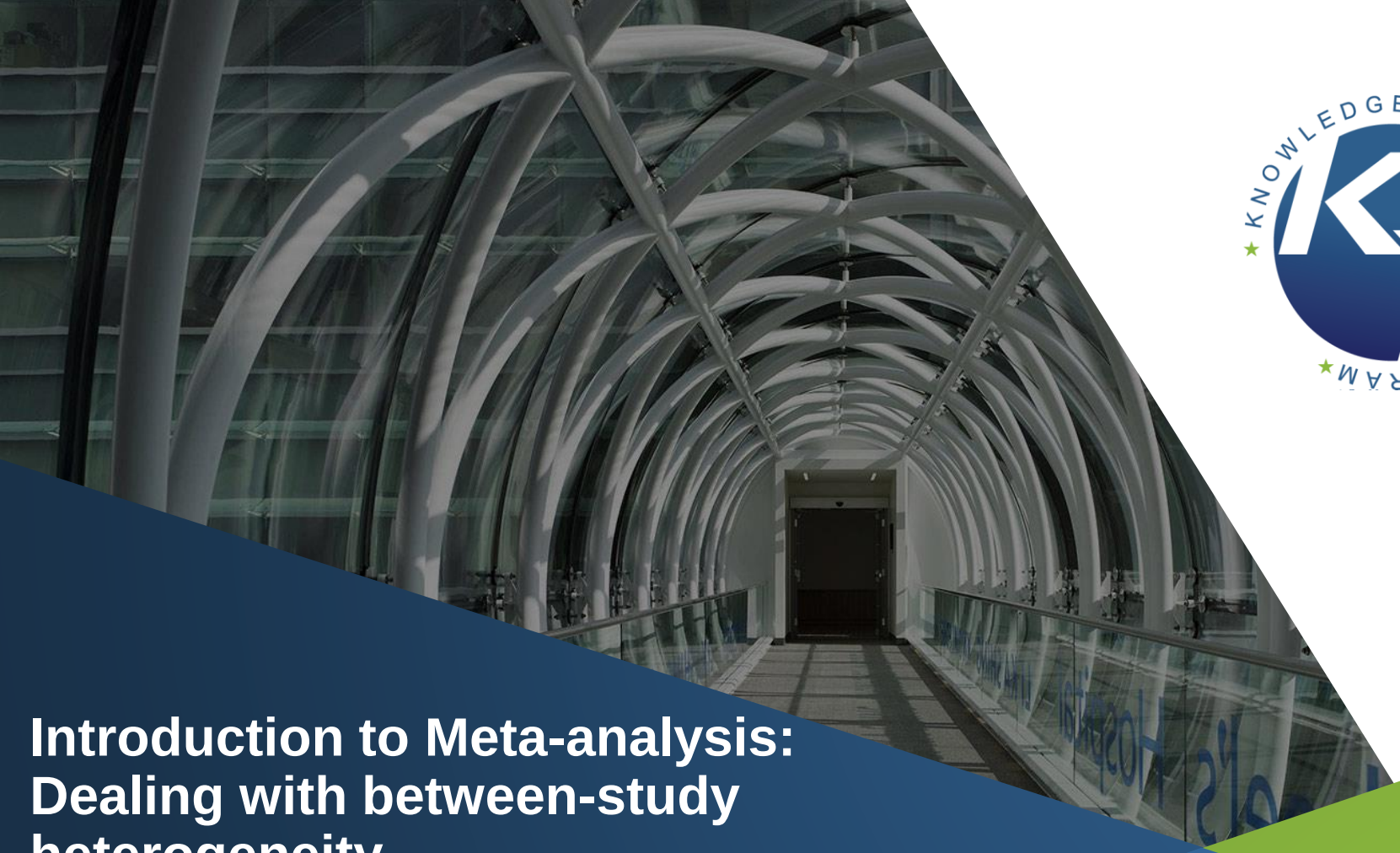


Introduction to Meta-analysis: Dealing with between-study heterogeneity

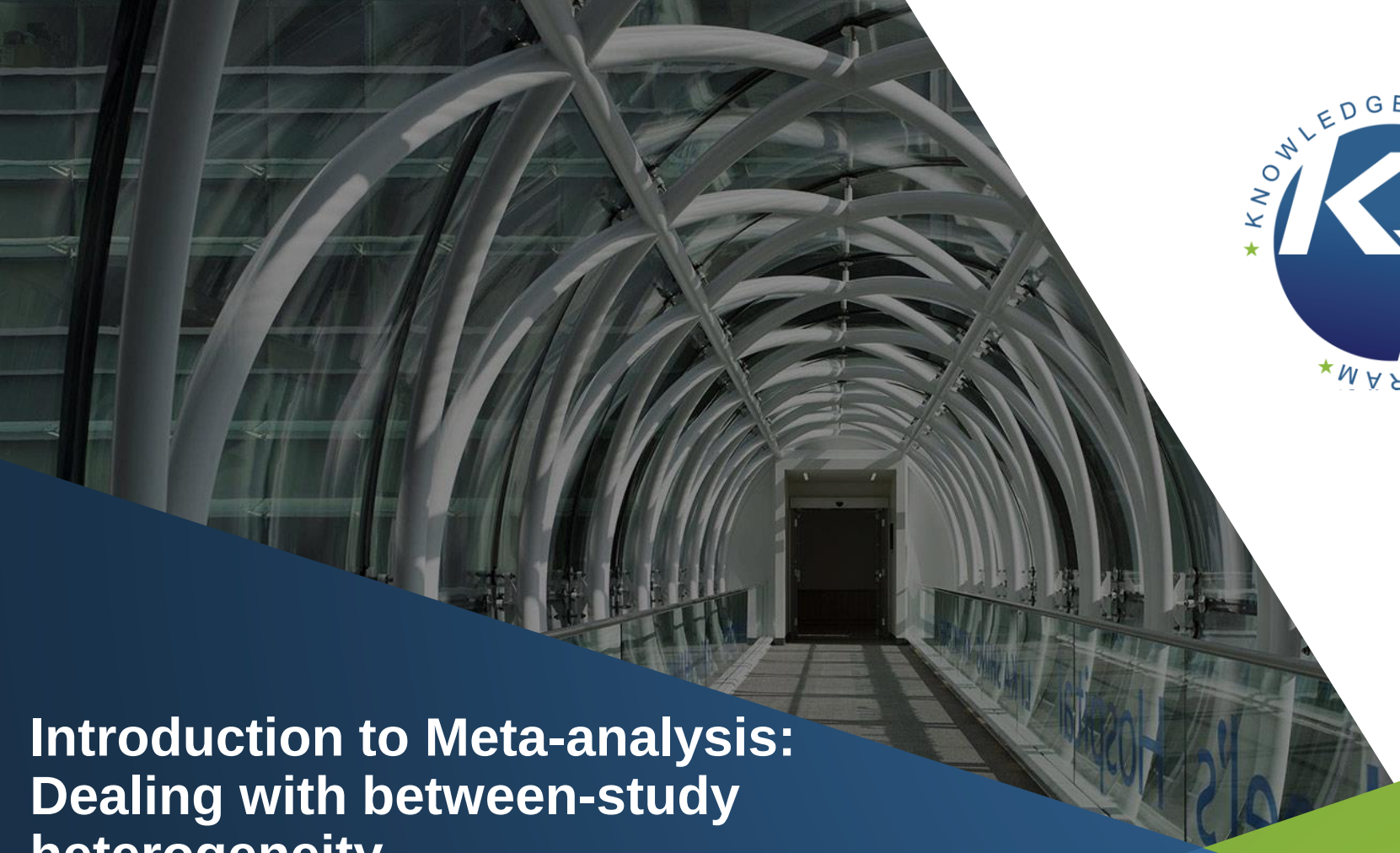
St. Michael's
Inspired Care.
Inspiring Science.

Dr. Areti Angeliki Veroniki

Scientist: Knowledge Translation Program, Li Ka Shing Knowledge Institute of St. Michael's Hospital

Speaker's Bio:

Dr. Areti-Angeliki Veroniki is a mathematician, holds an MSc in Statistics and Operations Research, and a PhD in Epidemiology. She is a Scientist at the Knowledge Translation Program of St. Michael's Hospital (appointed in July 2017), whose research interests are in optimizing the processes of evidence-based medicine. In particular, her research focuses on the statistical modelling for knowledge synthesis and the methodology of systematic and rapid reviews. She is interested in enhancing methods for meta-analysis and network meta-analysis with aggregated data and/or individual patient data, as well as in developing models to incorporate dosages and complex interventions in network meta-analysis. She works closely with the Cochrane Collaboration and the groups Comparing Multiple Interventions Methods Group (CMIMG), Individual Patients Data Meta-Analysis Methods Group (IPDMAG), and Statistical Methods Group (SMG), and she is a Statistical Editor for the Cochrane Depression, Anxiety and Neurosis Group (CDANG), and the Cochrane Developmental, Psychosocial and Learning Problems Group (CDPLPG). She is an Associate Editor for the international journal Systematic Reviews.



Introduction to Meta-analysis: Dealing with between-study heterogeneity

St. Michael's
Inspired Care.
Inspiring Science.

Dr. Areti Angeliki Veroniki

Scientist: Knowledge Translation Program, Li Ka Shing Knowledge Institute of St. Michael's Hospital

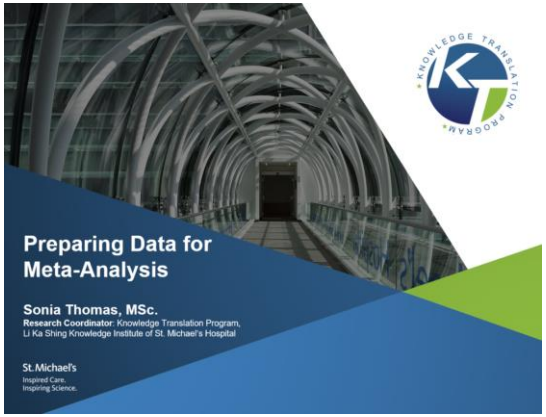
Learning objectives

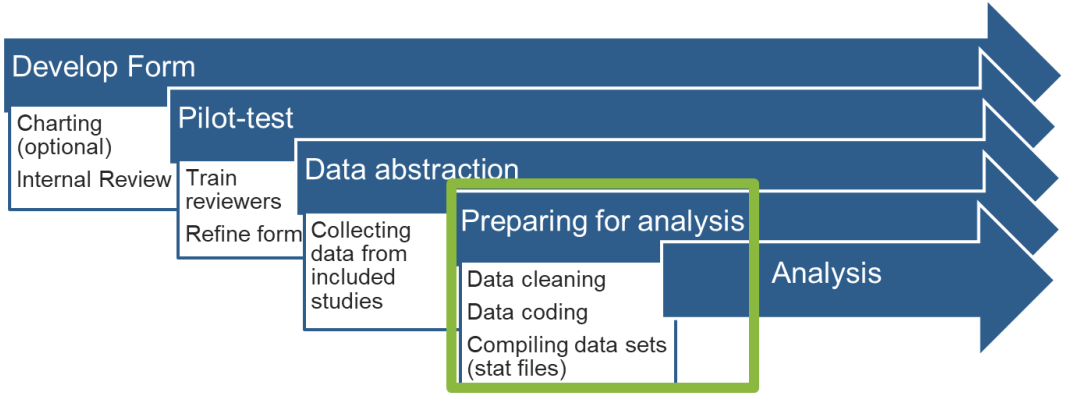
- Explain what between-study heterogeneity is

- Provide ways of
  - identifying,
  - dealing with, and
  - exploring between-study heterogeneity

St. Michael's
Inspired Care.
Inspiring Science.

To apply a meta-analysis



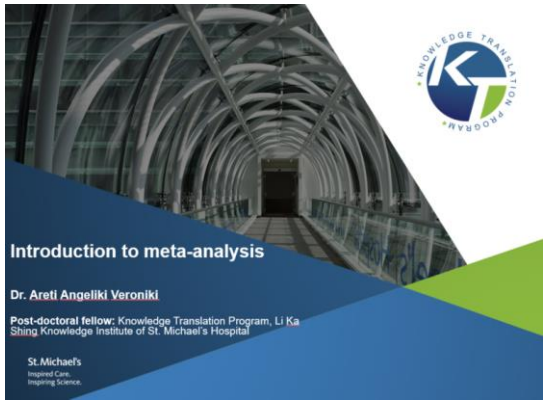


1. Identify the data type for the outcome measurements

2. Use an effect size to compare the outcomes between the interventions
   - E.g., Odds ratio, risk ratio, risk difference for binary data
   - E.g., Mean difference, standardized mean difference for continuous data

St. Michael's
Inspired Care.
Inspiring Science.

To apply a meta-analysis



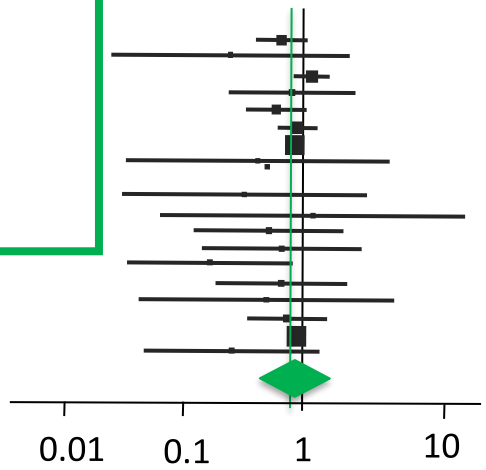
3. Extract data from each study:
   - estimate of treatment effect
   - variance of estimate

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

4. Combine these using a weighted average:

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

$$\textbf{with variance} = \frac{1}{\text{sum of weights}}$$



St. Michael's
Inspired Care.
Inspiring Science.

Group discussion

How many of you do you think we can add data from all studies together?

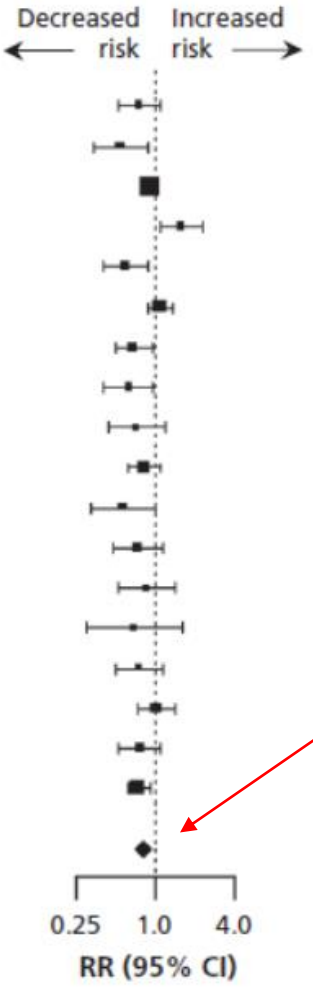
Can we add the treatment groups together and all the control groups together and compare the totals?

☐ YES ☐ NO

St. Michael's
Inspired Care.
Inspiring Science.

Can't we add up the data from all trials together?

Study	Treatment *n/N*	Control *n/N*	RR (95%CI)
Beck et al.[51]	35/160	47/161	0.75 (0.51 to 1.09)
Botha et al.[25]	13/32	18/24	0.54 (0.34 to 0.87)
Burns et al.[45]	210/353	228/355	0.93 (0.82 to 1.04)
Franklin et al.[60]	62/213	38/204	1.56 (1.10 to 2.23)
Lafave et al.[53]	13/24	37/41	0.60 (0.41 to 0.88)
Puschner et al.[29]	108/241	103/250	1.09 (0.89 to 1.33)
Rich et al.[55]	41/142	59/140	0.69 (0.50 to 0.95)
Salkever et al.[48]	27/91	25/53	0.63 (0.41 to 0.96)
Rich et al.[58]	21/63	16/35	0.73 (0.44 to 1.20)
Kasper et al.[43]	47/102	55/98	0.82 (0.62 to 1.08)
Courtney et al.[30]	13/49	27/58	0.57 (0.33 to 0.98)
Castro et al.[40]	20/50	25/46	0.74 (0.48 to 1.13)
Burns et al.[26]	17/110	56/313	0.86 (0.53 to 1.42)
Koehler et al.[32]	6/20	9/21	0.70 (0.30 to 1.61)
Ruchlewska et al.[28]	24/70	33/73	0.76 (0.50 to 1.14)
Laramee et al.[41]	49/131	46/125	1.02 (0.74 to 1.40)
Stewart et al.[50]	24/49	31/48	0.76 (0.53 to 1.08)
Lichtenberg et al.[34]	71/122	74/95	0.75 (0.62 to 0.90)
Overall Heterogeneity: I^2 = 58%	**801/2022**	**927/2140**	0.81 (0.72 to 0.91)



Tricco et al (2014)

The RR calculated in a meta-analysis is RR=0.81 – decreases the risk of being admitted to the hospital by 19%

If we add up the columns we get a risk of 0.40 vs. 0.43, i.e. a RR of 0.91 – decreases the risk of being admitted to the hospital by 9%

St. Michael's
Inspired Care.
Inspiring Science.

Why can't we add up the data from all trials together?

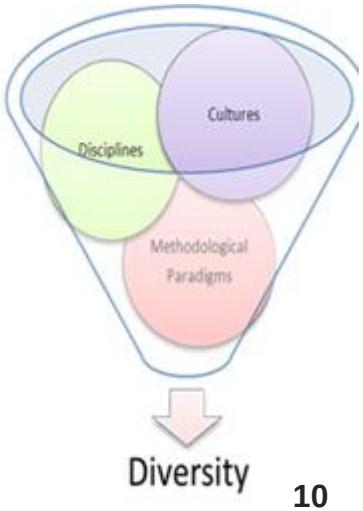
- Breaks within-study randomization
- Assumes that all patients belong to a single mega-trial
- Imbalances within-trials can introduce bias
- Does not account for potential between-study heterogeneity
- Should be avoided!

Between-study Heterogeneity

- Rapid reviews usually include studies with different
  - People
  - Countries
  - Length of time
  - Settings
  - Outcomes
  - …..

- Heterogeneity refers to the differences observed between studies included in a rapid review

- Types of heterogeneity:
  1. Clinical
  2. Methodological
  3. Statistical



St. Michael's
Inspired Care.
Inspiring Science.

Clinical Between-study Heterogeneity

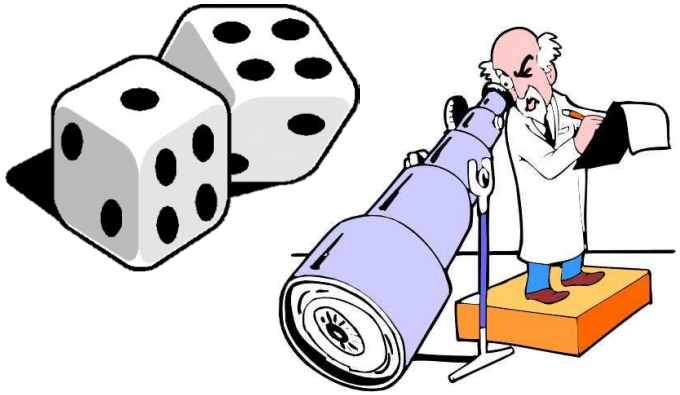
- *Patients*
  - e.g. inclusion and exclusion criteria for trials, diagnosis, geographical variation

- *Interventions*
  - e.g. intensity, dose, duration, type of drug, mode of administration, nature of the control (placebo/none/standard care)

- *Outcomes*
  - e.g. type, cut-off points, definition of an event, follow-up duration, ways of measuring outcomes, scale

Methodological Between-study Heterogeneity

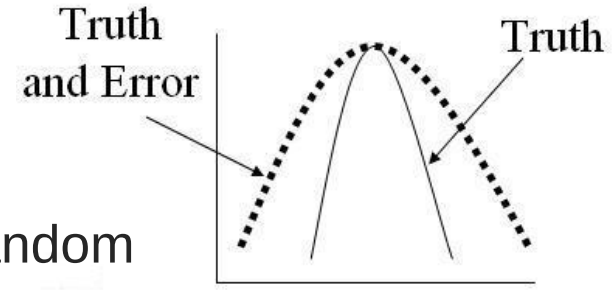
- *Design*
  - e.g., RCTs vs. non-randomized studies,
    crossover vs. parallel group vs. cluster randomised, length of study



- *Conduct*
  - e.g., allocation concealment, blinding, analysis approach, imputation approaches for missing data

St. Michael's
Inspired Care.
Inspiring Science.

Statistical Between-study Heterogeneity



- *Commonly considered*
  - Diversity in the study results
  - More variation than would be expected by random error

- *In truth:*
  - Variation in the true effects underlying the studies
  - When homogeneity does not hold

  (homogeneity = identical effect underlying every study)

How to look for between-study heterogeneity?

1. Visually
   - inspection of the forest plots: do study confidence intervals overlap?

2. Statistically
   - Apply tests for heterogeneity (e.g. Cochran's Q-test, Generalized Q-test)
     - The Cochran's Q-test (or chi-square test for heterogeneity) tests whether the study-specific effects differ from the meta-analysis effect beyond what is expected by random error
   - Use statistics to infer on the between-study variability: I^2, H^2, R^2, D^2, G^2
     - I^2: the percentage of total variability across studies that is due to between-study heterogeneity beyond what is expected by random error
       a. Based on Cochran's homogeneity statistic (Higgins and Thompson 2002)
       b. Based on the Generalized Q-statistic (Bowden et al 2011)
   - Estimate the magnitude of the between-study variance (τ^2)

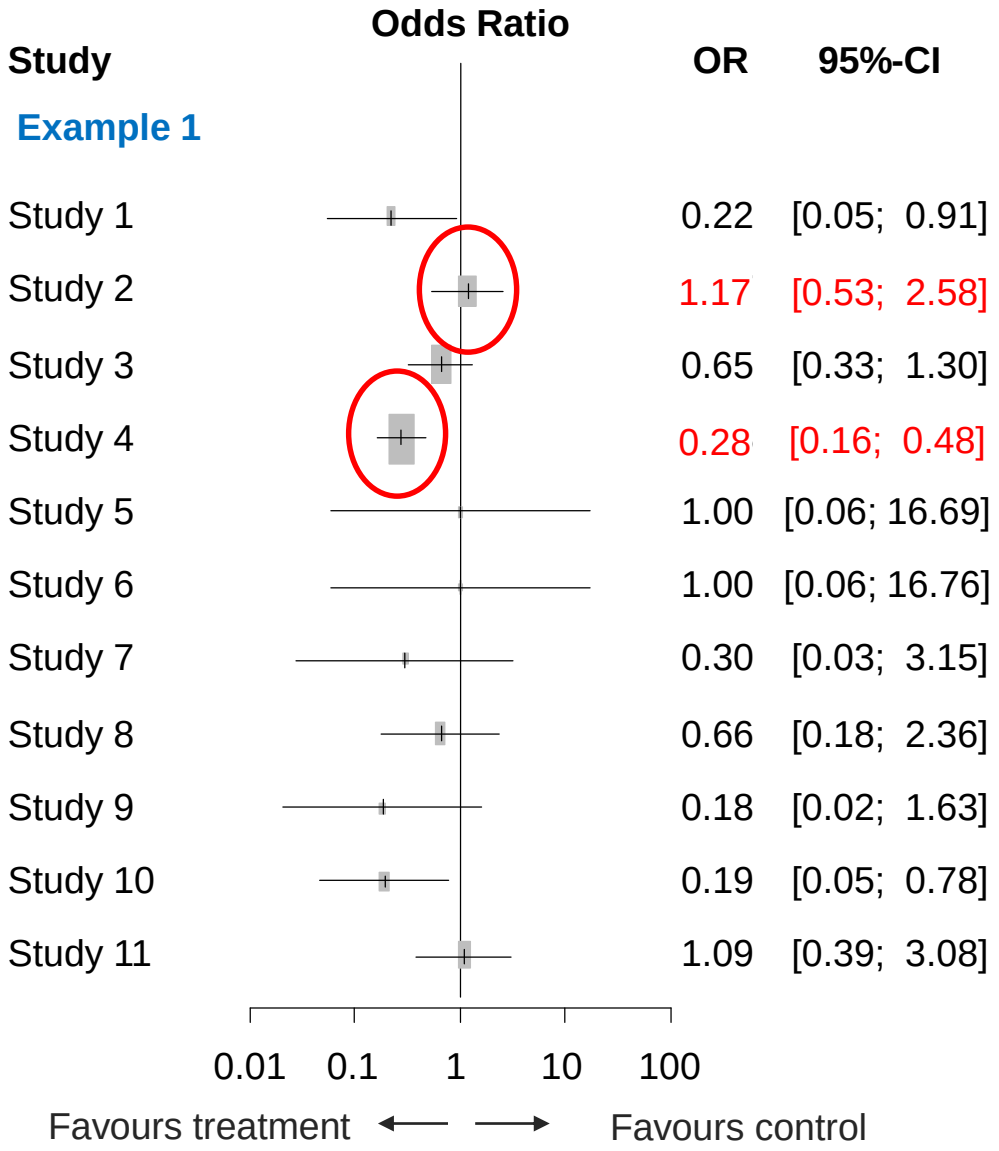
Group discussion

In the following two examples, is there a between-study variability?

Identifying between-study heterogeneity - Visually

☐ YES

☐ NO

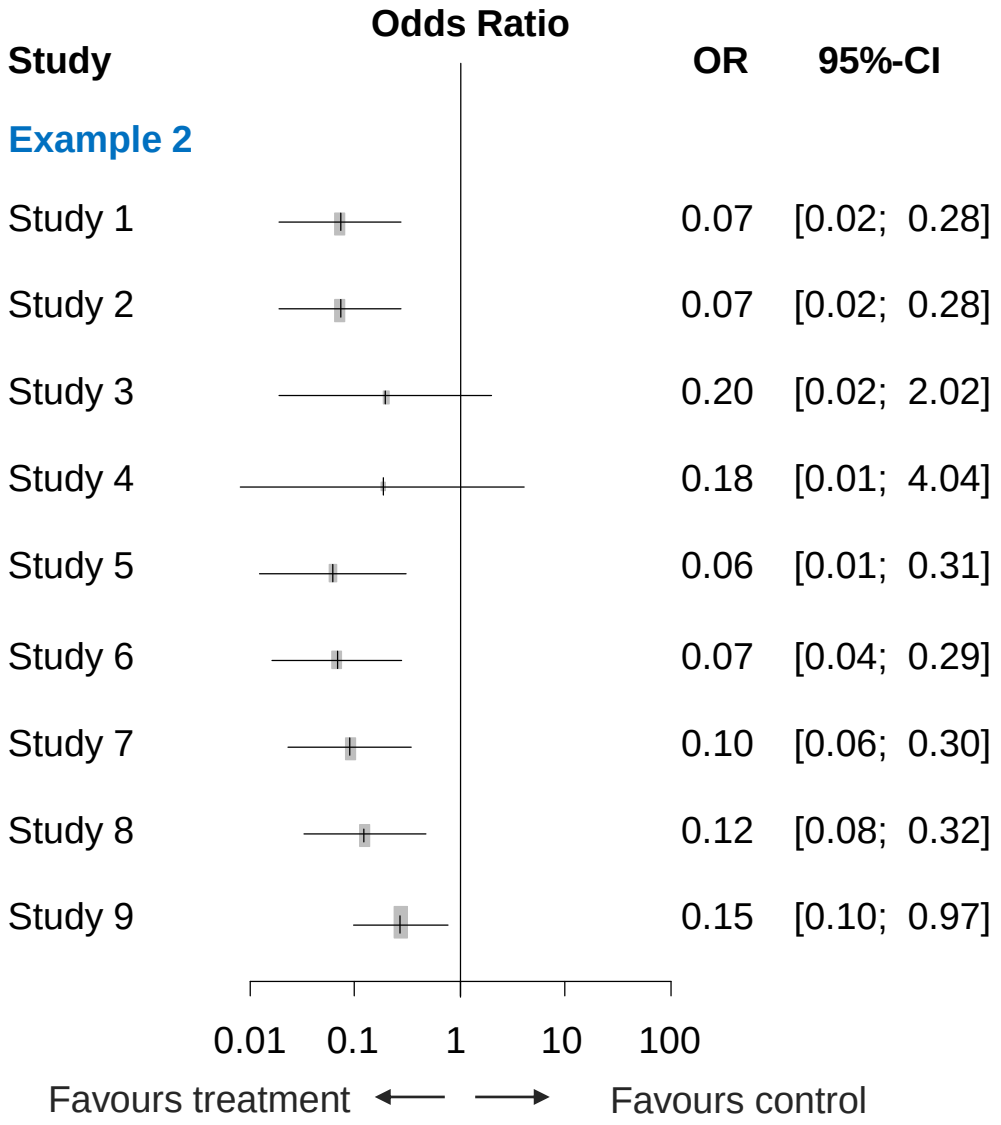


St. Michael's
Inspired Care.
Inspiring Science.

Identifying between-study heterogeneity - Visually

☐ YES

☐ NO



St. Michael's
Inspired Care.
Inspiring Science.

Identifying between-study heterogeneity - Statistically

Q-test

Under the assumption that $\tau^2 = 0$, the Cochran's Q-statistic (or chi-squared (χ^2) test) is:

$$Q = \sum_{i=1}^{k} w_{i,FE}(y_i - \hat{\mu}_{FE})^2$$

- Has χ^2 distribution with k – 1 d.f. under null hypothesis
- Larger values of Q reflect greater between-study heterogeneity
- Rejection of H_0: $\tau^2 = 0$ suggests heterogeneity

w_i: weight in study i

y_i: effect size in study i

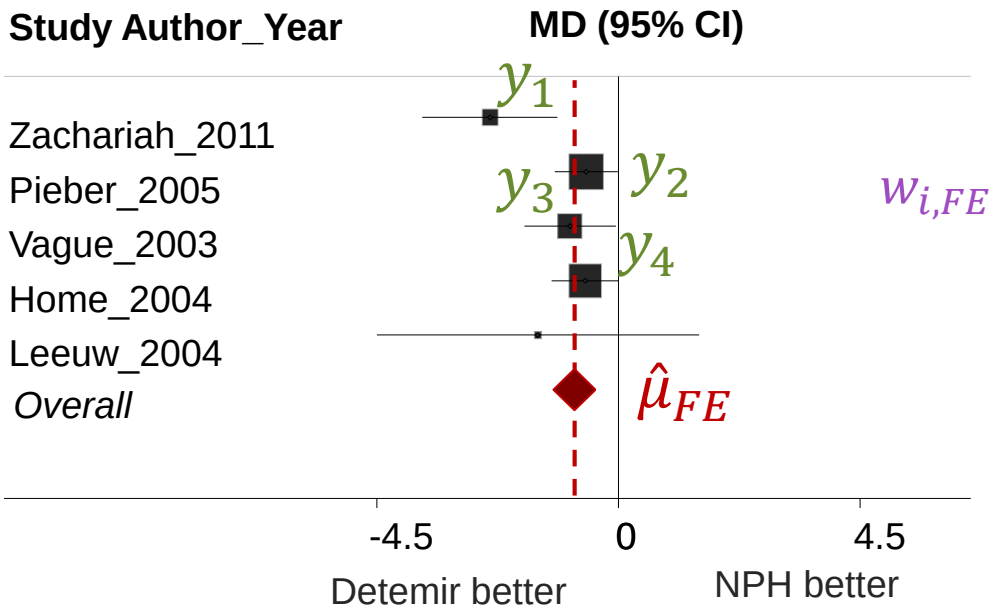
μ: pooled estimate

k: number of studies in meta-analysis

Identifying between-study heterogeneity - Statistically

Cochran's Q-statistic (or chi-squared (χ^2) test) - Graphical illustration

Detemir vs. NPH for preventing weight gain in patients with type 1 diabetes



$$w_{i,FE} = \frac{1}{Var(y_i)}$$

$$Q = \sum_{i=1}^{k} w_{i,FE}(y_i - \hat{\mu}_{FE})^2$$

w_i: weight in study i

y_i: effect size in study i

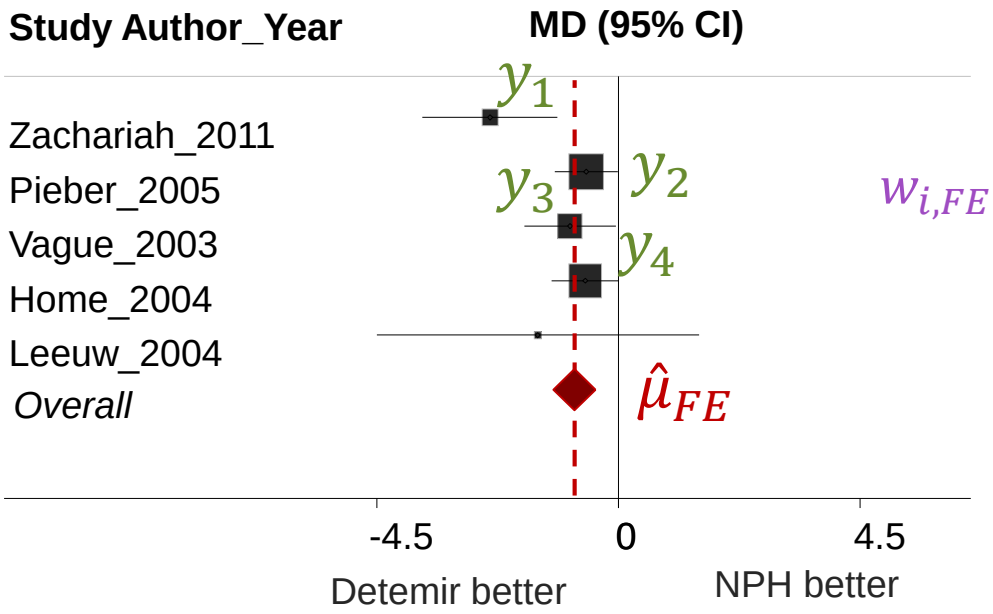
μ: pooled estimate

k: number of studies in meta-analysis

Identifying between-study heterogeneity - Statistically

Cochran's Q-statistic (or chi-squared (χ^2) test) - Graphical illustration

Detemir vs. NPH for preventing weight gain in patients with type 1 diabetes



$$w_{i,FE} = \frac{1}{Var(y_i)}$$

$$Q = 7.15\ (d.f. = 4)$$
$$p = 0.128$$

Identifying between-study heterogeneity - Statistically

Cochran's Q-statistic (or chi-squared (χ^2) test) - Graphical illustration

Detemir vs. NPH for preventing weight gain in patients with type 1 diabetes

Does the chi-square test suggest that there is between-study heterogeneity in this meta-analysis?

☐ YES ☐ NO

$$Q = 7.15\ (d.f. = 4)$$
$$p = 0.128$$

Identifying between-study heterogeneity - Statistically

Cochran's Q-statistic (or chi-squared (χ^2) test)

- When the number of studies is small, Cochran's Q test has low power (cut-off P-value: 0.10; i.e., P-value<0.10 indicates significant between-study heterogeneity)

- If there are many studies included, small heterogeneity may be found to be statistically significant.

Hardy and Thompson (1998)

Generalized Q-test

Under the assumption that $\tau^2 = \tau_0^2$ ($\tau_0^2 \geq 0$) the generalized Q-statistic:

$$Q_{gen} = \sum_{i=1}^{k} w_{i,RE}(y_i - \hat{\mu}_{RE})^2$$

w_i: weight in study i

y_i: effect size in study i

μ: pooled estimate

k: number of studies in meta-analysis

St. Michael's
Inspired Care.
Inspiring Science.

Identifying between-study heterogeneity - **Statistically**

Cochran's Q-statistic (or chi-squared (χ^2) test)

Assumes that all studies share a common true effect size

Generalized Q-test

Assumes that each study has each own true effect size

Identifying between-study heterogeneity - **Statistically**

I^2 statistic

Describes the proportion of variability that is due to heterogeneity rather than sampling error

I^2 - based on the Cochran's homogeneity statistic

- Quantifies heterogeneity based on the Cochran's homogeneity statistic and its degrees of freedom

$$I^2 = \frac{Q-(k-1)}{Q} \cdot 100\%$$

Higgins and Thompson 2002

St. Michael's
Inspired Care.
Inspiring Science.

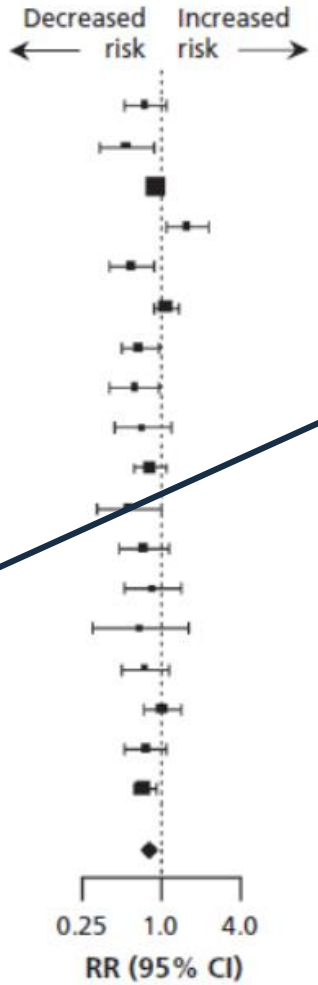
Group discussion

In the following example, what is the proportion of the observed variance that is due to real differences in effect sizes? Can you identify this in the forest plot?

Identifying between-study heterogeneity - Statistically

I^2 - based on the Cochran's homogeneity statistic

Study	Treatment n/N	Control n/N	RR (95%CI)
Beck et al.[51]	35/160	47/161	0.75 (0.51 to 1.09)
Botha et al.[25]	13/32	18/24	0.54 (0.34 to 0.87)
Burns et al.[45]	210/353	228/355	0.93 (0.82 to 1.04)
Franklin et al.[60]	62/213	38/204	1.56 (1.10 to 2.23)
Lafave et al.[53]	13/24	37/41	0.60 (0.41 to 0.88)
Puschner et al.[29]	108/241	103/250	1.09 (0.89 to 1.33)
Rich et al.[55]	41/142	59/140	0.69 (0.50 to 0.95)
Salkever et al.[48]	27/91	25/53	0.63 (0.41 to 0.96)
Rich et al.[58]	21/63	16/35	0.73 (0.44 to 1.20)
Kasper et al.[43]	47/102	55/98	0.82 (0.62 to 1.08)
Courtney et al.[30]	13/49	27/58	0.57 (0.33 to 0.98)
Castro et al.[40]	20/50	25/46	0.74 (0.48 to 1.13)
Burns et al.[26]	17/110	56/313	0.86 (0.53 to 1.42)
Koehler et al.[32]	6/20	9/21	0.70 (0.30 to 1.61)
Ruchlewska et al.[28]	24/70	33/73	0.76 (0.50 to 1.14)
Laramee et al.[41]	49/131	46/125	1.02 (0.74 to 1.40)
Stewart et al.[50]	24/49	31/48	0.76 (0.53 to 1.08)
Lichtenberg et al.[34]	71/122	74/95	0.75 (0.62 to 0.90)
Overall			0.81 (0.72 to 0.91)
Heterogeneity: I^2 = 58%			



Tricco et al (2014)

$I^2 = 58\%$

Identifying between-study heterogeneity - Statistically

I^2 statistic

I^2 - based on the Cochran's homogeneity statistic

- *D*epends on the number (for k<10) and size of studies - increases as size increases
- I^2 is surrounded by a magnitude of uncertainty – a CI would help infer on its uncertainty
- CIs for I^2 provide good coverage as evidence accumulates

I^2 - based on the Generalized Q-statistic

Bowden et al 2011

- Maintains well the desired coverage compared to I^2 based on the Cochran's homogeneity statistic
- Wider CIs for I^2 than those of I^2 using Cochran's Q

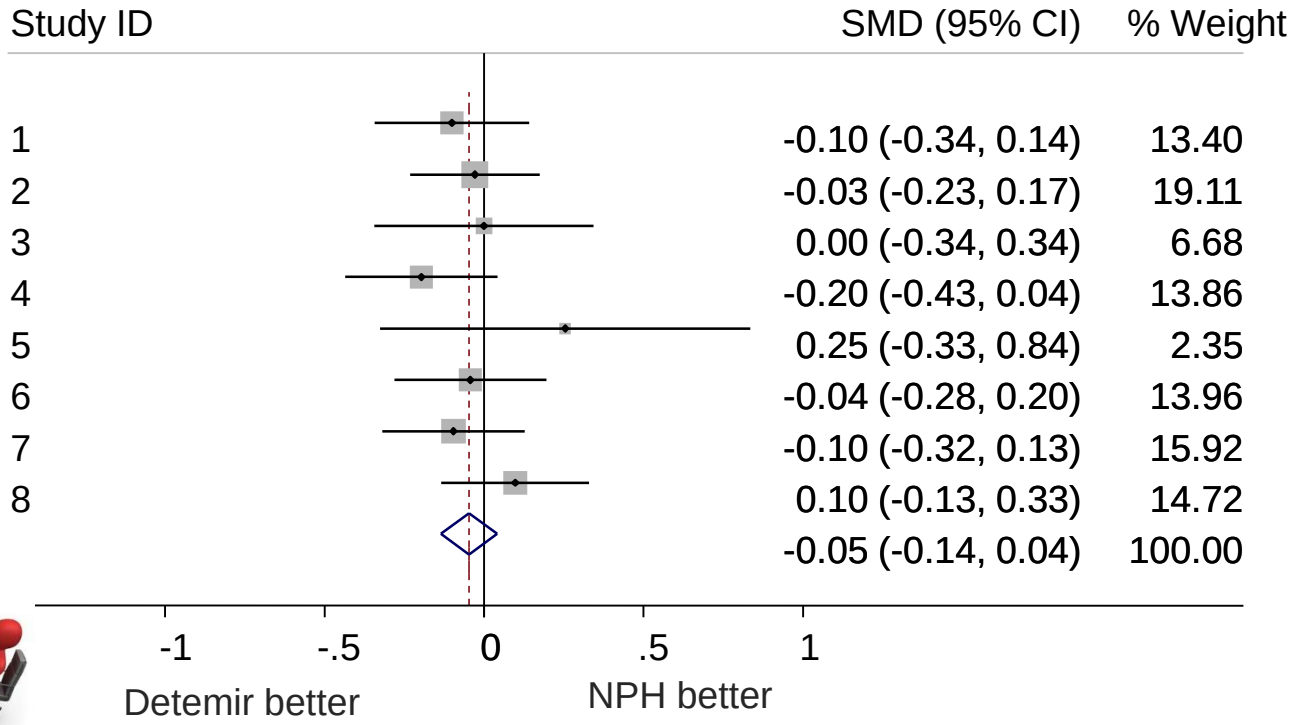
Identifying between-study heterogeneity - **Statistically**

Between-study variance (τ^2)

- The most popular approach is the DerSimonian and Laird approach to estimate the between-study variance
- There are 16 different approaches to estimate the true τ^2 (for a review of methods see Veroniki et al 2016)
- The use of CIs for τ^2 can help in the interpretation of the between-study variance

Identifying between-study heterogeneity - Statistically

Detemir vs. NPH for A1C in patients with type 1 diabetes



$$Q = 4.24\,(d.f. = 7)\,,\ p = 0.751$$
$$I^2 = 0\%$$
$$\tau^2 = 0.00$$

St. Michael's
Inspired Care.
Inspiring Science.

Identifying between-study heterogeneity - Statistically

Detemir vs. NPH for A1C in patients with type 1 diabetes

Are the study-specific effects similar enough that we are confident a combined estimate will be a meaningful result?

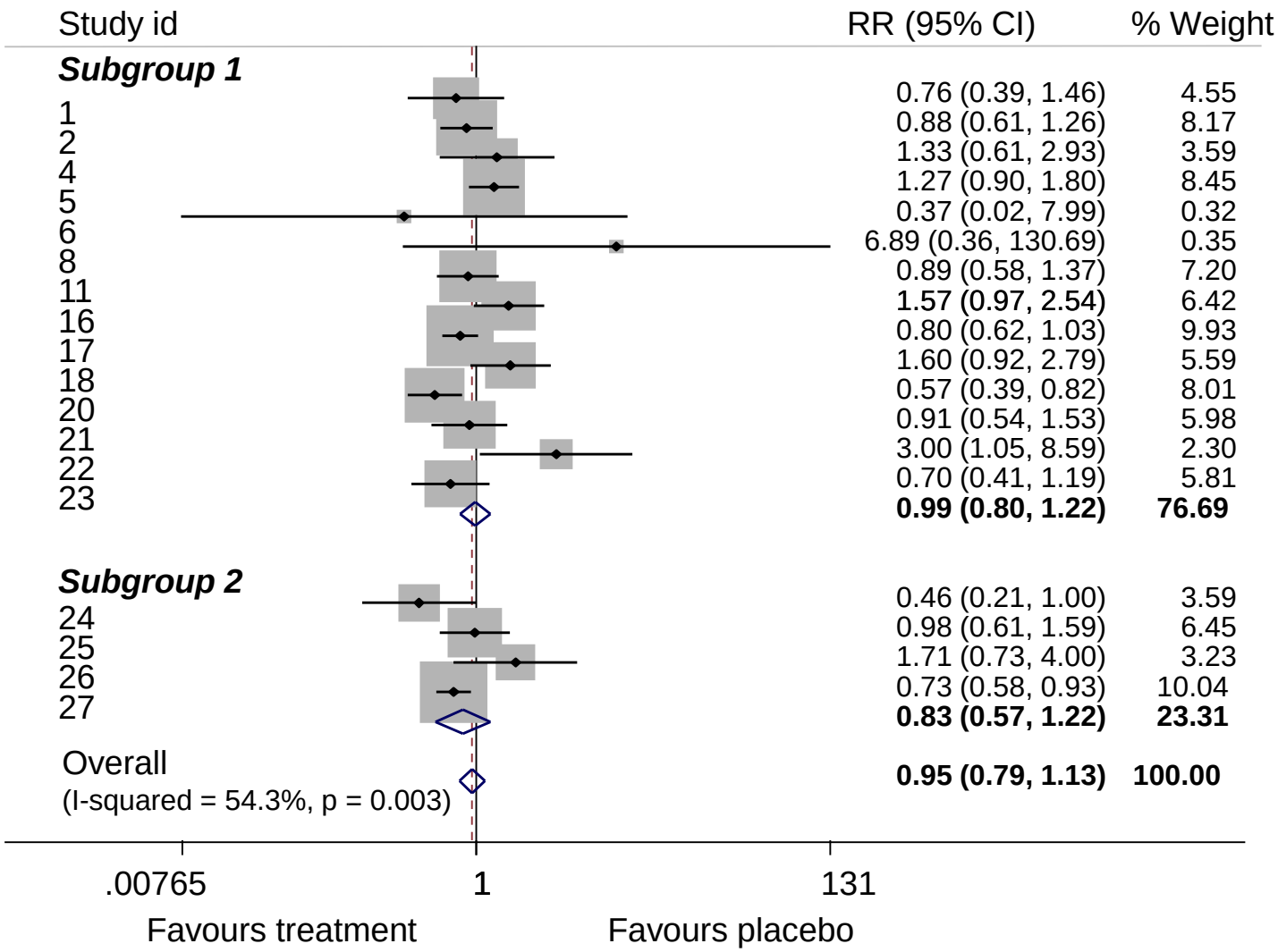
☐ YES ☐ NO

$$Q = 4.24\ (d.f. = 7)\ ,\ p = 0.751$$
$$I^2 = 0\%$$
$$\tau^2 = 0.00$$

St. Michael's
Inspired Care.
Inspiring Science.

Identifying between-study heterogeneity - Statistically

Fictional example

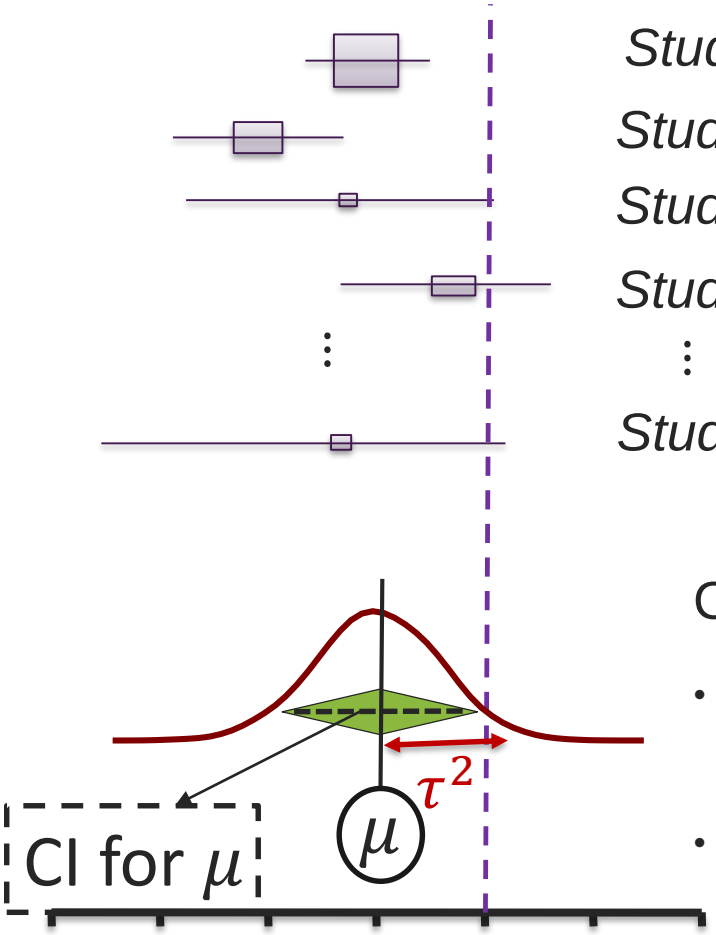


Identifying between-study heterogeneity - Statistically

Fictional example

Subgroup	*Chi-square statistic*	*Degrees of freedom*	*P-value*	*I-square*	*Tau-square*
1	28.96	13	0.007	55.1%	0.0734
2	6.23	3	0.101	51.8%	0.0732
Overall	37.20	17	0.003	54.3%	0.0660

Meta-analysis



Study 1 with treatment effect	y_1 *and variance*	v_1
Study 2	y_2	v_2
Study 3	y_3	v_3
Study 4	y_4	v_4
⋮		
Study k	y_k	v_k

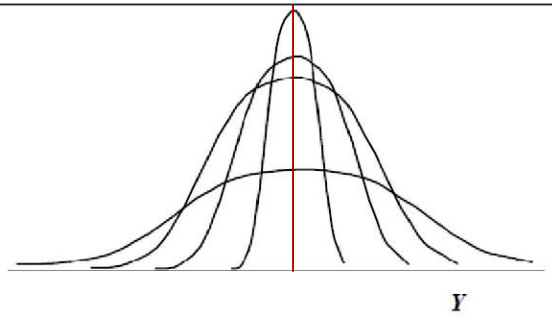
Considerations:

- Heterogeneity suggests that the studies have important underlying differences
- We assume the true study-specific effects follow a distribution (usually a normal distribution)

St. Michael's
Inspired Care.
Inspiring Science.

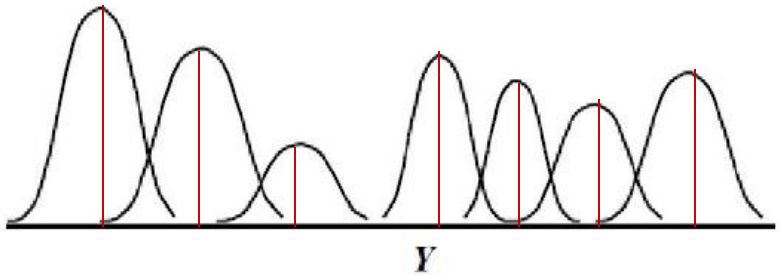
Meta-analysis models

<u>Fixed-effect model assumption</u>:
Studies are sufficiently *similar* in aspects that could modify the treatment effect



There is a single true effect for all studies

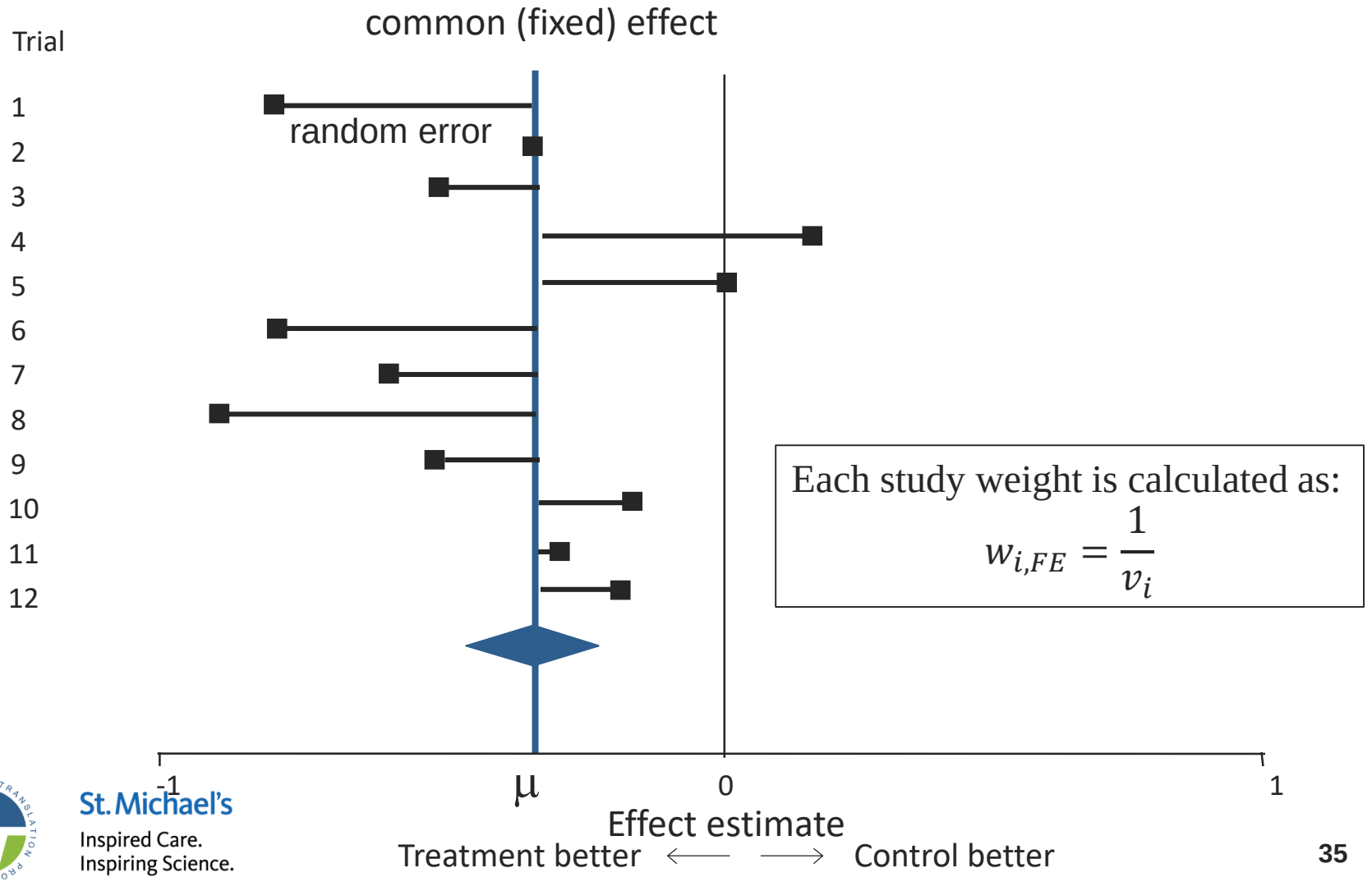
<u>Random-effects model assumption</u>:
The observed study-specific effects estimate different *true* effects, which are related and come from the same distribution



Each study has each own true effect – all study-specific true effects are exchangeable

St. Michael's
Inspired Care.
Inspiring Science.

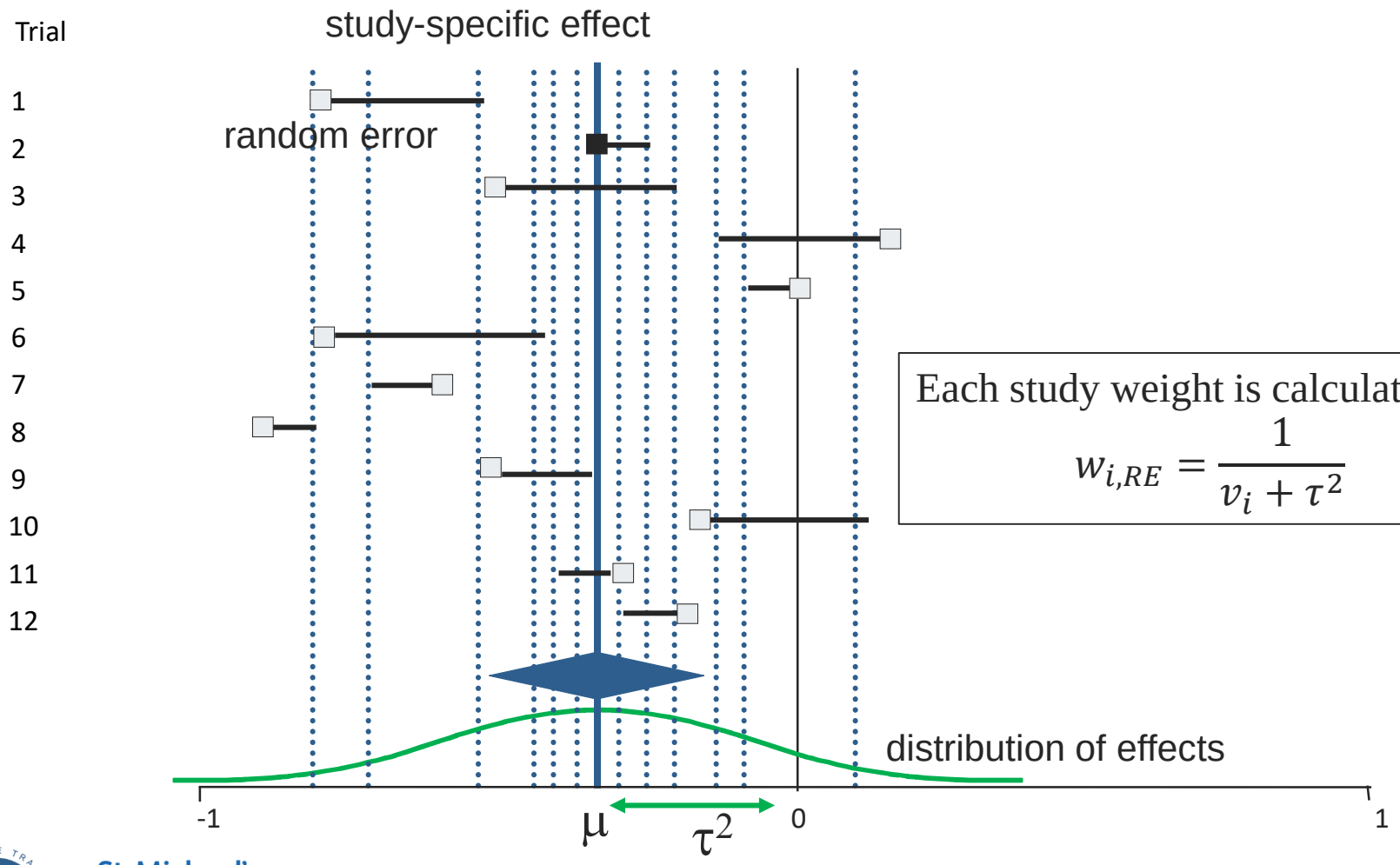
Fixed-effect meta-analysis



Each study weight is calculated as:

$$w_{i,FE} = \frac{1}{v_i}$$

Random-effects meta-analysis



Each study weight is calculated as:

$$w_{i,RE} = \frac{1}{v_i + \tau^2}$$

St. Michael's
Inspired Care.
Inspiring Science.

Fixed-effect (FE) vs. Random-effects (RE) model

- Fixed and random-effects meta-analyses may be identical (for $\tau^2 = 0$)
- Fixed-effect model often unrealistic
  - "Since systematic reviews bring together studies that are diverse both clinically and methodologically, heterogeneity in their results is to be expected."

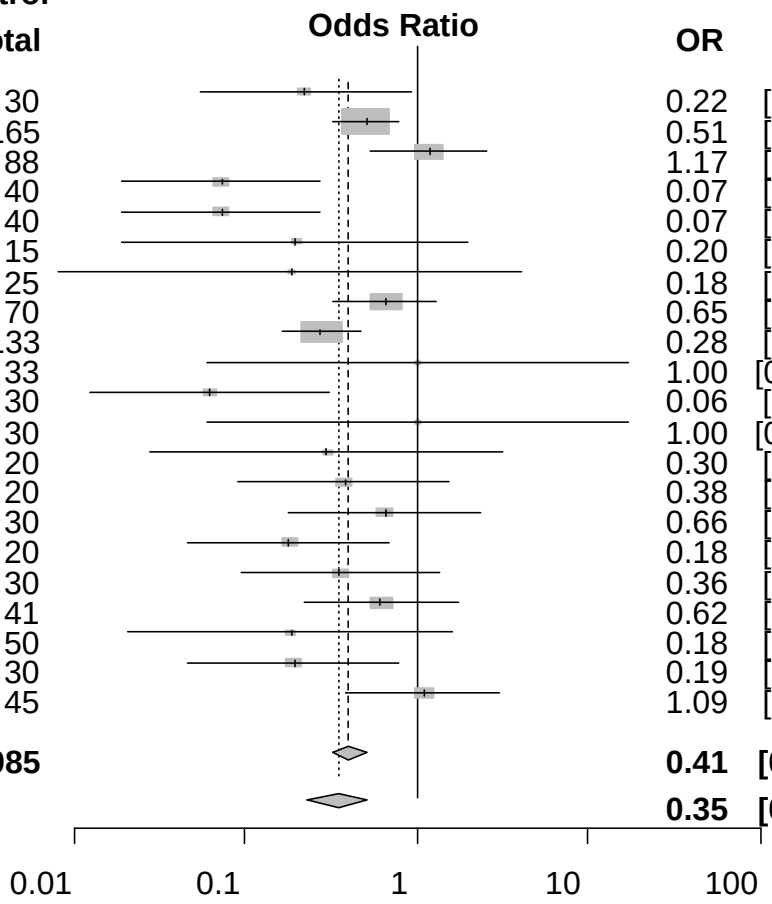
Higgins et al., BMJ 2003

- Random-effects model is difficult to justify
- Random-effects analysis may give spurious results when effect size depends on precision
  - Gives relatively more weight to smaller studies
  - Smaller studies may be of lower quality (hence biased)
  - Publication bias may result in missing smaller studies (or small-study effects; FE and RE overall effects might disagree)
- Random-effects meta-analysis suitable for unexplained small to moderate heterogeneity

Fixed-effect (FE) vs. Random-effects (RE) model

- The random-effects model gives more conservative results

Study	Experimental Events	Experimental Total	Control Events	Control Total	OR	95%-CI	Weight (fixed)	Weight (random)
119718	3	30	10	30	0.22	[0.05; 0.91]	2.5%	4.5%
119782	58	165	85	165	0.51	[0.33; 0.79]	25.7%	9.3%
119985	16	88	14	88	1.17	[0.53; 2.58]	8.1%	7.5%
124097	19	40	37	40	0.07	[0.02; 0.28]	2.8%	4.9%
124098	19	40	37	40	0.07	[0.02; 0.28]	2.8%	4.9%
124295	1	15	4	15	0.20	[0.02; 2.02]	0.9%	2.3%
125714	0	25	2	25	0.18	[0.01; 4.04]	0.5%	1.4%
126278	40	70	47	70	0.65	[0.33; 1.30]	10.6%	8.0%
126721	67	134	104	133	0.28	[0.16; 0.48]	17.7%	8.9%
R59	1	33	1	33	1.00	[0.06; 16.69]	0.6%	1.7%
120183	2	30	16	30	0.06	[0.01; 0.31]	2.0%	3.9%
120185	1	30	1	30	1.00	[0.06; 16.76]	0.6%	1.7%
121278	1	20	3	20	0.30	[0.03; 3.15]	0.9%	2.3%
121423	4	20	8	20	0.38	[0.09; 1.54]	2.5%	4.5%
122761	5	30	7	30	0.66	[0.18; 2.36]	3.1%	5.1%
123056	5	20	13	20	0.18	[0.05; 0.70]	2.7%	4.7%
123155	4	30	9	30	0.36	[0.10; 1.33]	2.9%	4.9%
123482	8	36	13	41	0.62	[0.22; 1.71]	4.8%	6.2%
124081	1	50	5	50	0.18	[0.02; 1.63]	1.1%	2.5%
124213	3	30	11	30	0.19	[0.05; 0.78]	2.5%	4.6%
S2	9	42	9	45	1.09	[0.39; 3.08]	4.7%	6.2%
Fixed effect model	**978**		**985**		**0.41**	**[0.33; 0.51]**	**100.0%**	**--**
Random effects model					**0.35**	**[0.23; 0.52]**	**--**	**100.0%**



St. Michael's
Inspired Care.
Inspiring Science.

What can we do with heterogeneity?

Resign to it	Do not do pool data into a meta-analysis
Ignore it	Use a fixed-effect model
Allow for it	Use a random-effects model
Explore it	• Check the data for incorrect data extraction; unit of analysis errors (e.g. with crossover trials) • Change effect measure: Ratio measures (RR and OR) considerably less heterogeneous than difference measure (RD) • Subgroup analysis • Meta-regression

St. Michael's
Inspired Care.
Inspiring Science.

Which of the following approaches have you ever used?

Resign to it	Do not do pool data into a meta-analysis
Ignore it	Use a fixed-effect model
Allow for it	Use a random-effects model
Explore it	• Check the data for incorrect data extraction; unit of analysis errors (e.g. with crossover trials) • Change effect measure: Ratio measures (RR and OR) considerably less heterogeneous than difference measure (RD) • Subgroup analysis • Meta-regression

St. Michael's
Inspired Care.
Inspiring Science.

What we should *NOT* do with heterogeneity

- Fixed-effect or random-effects meta-analysis should be specified *a priori* if possible and not on the basis of a heterogeneity test

What to do:

- Think about the question you aim to respond, the inclusion criteria and eligible studies - do you expect them to vary considerably?

- You can apply and present both fixed and random-effects

Example: "We decided to apply a RE model, as we expected methodological and clinical heterogeneity across the included studies"

Tricco et al., BMC Med 2015

Cochrane Handbook (section 9.5.2)

"Some argue that, since clinical and methodological diversity always occur in a meta-analysis, statistical heterogeneity is inevitable (Higgins 2003). Thus the test for heterogeneity is irrelevant to the choice of analysis; heterogeneity will always exist whether or not we happen to be able to detect it using a statistical test."

Group discussion

In the following example, which meta-analysis model should we use?

Which meta-analysis model should we use?

☐ Fixed-effect ☐ Random-effects

Aim of study:

To examine the influenza vaccine efficacy when this is not well matched to circulating strains

Study Protocol - PICOS:

"Healthy children, adults or older participants were chosen as our population of interest…

All influenza vaccines will be included…

will be limited to randomized clinical trials (RCTs) and quasi-RCTs comparing influenza vaccine(s) with placebo…

The primary outcome is the incidence of laboratory confirmed influenza identified by PCR or viral culture….

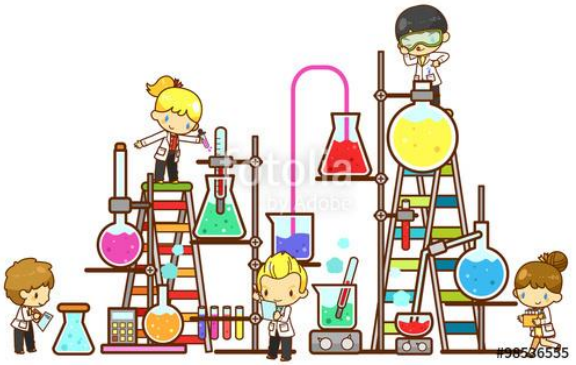
The secondary outcomes are laboratory-confirmed influenza identified by 1) antibody assay or 2) influenza infection determined by antibody assay, PCR, and/or viral culture"

Tricco et al (2013)

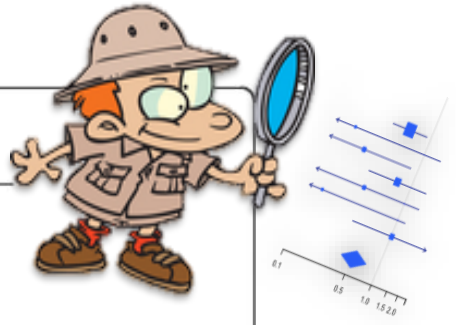
Which meta-analysis model should we use?

Table 2 Patient characteristics

Lead author (year)	Country of conduct and year	Age category	Mean age (SD) in years	M/F, %
Leibovitz (1971) [24]	USA, 1970	Adults	NR	NR
Beutner (1979) [25]	USA, 1974	Children	Range: 7 to 14	50/50
Rytel (1977) [26]	USA, 1974	Adults	NR	0/100
Monto (1982) [27]	USA, 1979	Adults	NR	NR
Tannock (1984) [28]	Australia, 1981	Adults	34.8 (13.9)	69/31
Keitel (1997) [29]	USA, 1983 to 1988	Adults	Range: 30 to 60	NR
Gruber (1990) [30]	USA, 1985	Children	7.9 (3.3)	NR
Edwards (1994) [31]	USA, 1986 to 1990	Adults/Children	Range: 1 to 65	NR
Clover (1991) [32]	USA, 1989	Children	8.8 (3.6)	NR
Govaert (1994) [33]	The Netherlands, 1991	Older patients	Range: 60 to 91	47/53
Powers (1995) [34]	USA, 1993	Adults	Range: 18 to 45	NR
Belshe (1998) [35]	USA, 1996	Children	3.5 (1.4)	47/53



Which meta-analysis model should we use?



Author(s) and Year	Vaccinated AB	Vaccinated Total	Control AB	Control Total	Relative Risk [95% CI]
Frey, 2010	30	3776	74	3843	0.41 [0.27 , 0.63]
Monto, 2009	1	813	5	325	0.08 [0.01 , 0.68]
Ohmit, 2008	0	853	1	338	0.13 [0.01 , 3.24]
Ohmit, 2006	10	519	10	206	0.40 [0.17 , 0.94]
Edwards, 1994	6	872	28	878	0.22 [0.09 , 0.52]
Edwards, 1994	12	1029	29	1064	0.43 [0.22 , 0.83]
Tam, 2007	54	1653	86	1111	0.42 [0.30 , 0.59]
Tam, 2007	7	503	16	494	0.43 [0.18 , 1.04]
Vesikari, 2006	11	951	38	665	0.20 [0.10 , 0.39]
Vesikari, 2006	7	640	9	450	0.55 [0.21 , 1.46]
BraccoNeto, 2009	3	944	2	942	1.50 [0.25 , 8.94]
BraccoNeto, 2009	23	338	32	342	0.73 [0.43 , 1.22]
Belshe, 2000	15	917	51	441	0.14 [0.08 , 0.25]
Forrest, 2008	23	525	24	516	0.94 [0.54 , 1.65]
Lum, 2010	14	765	11	385	0.64 [0.29 , 1.40]
RE Model					0.40 [0.29 , 0.56]

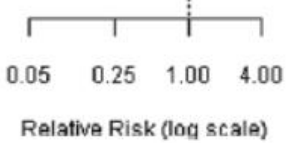


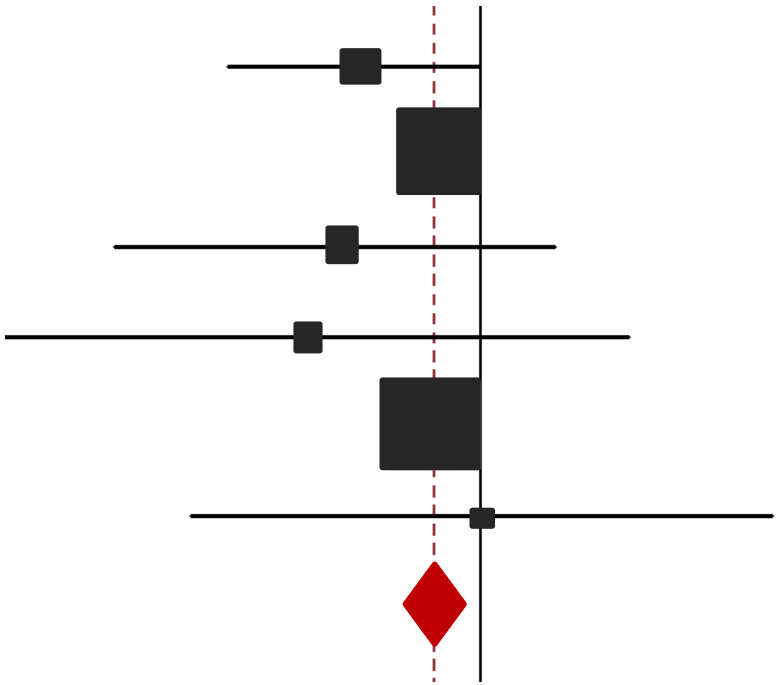
Figure 3 Meta-analysis of live attenuated influenza vaccine (LAIV), mismatched. This figure represents the relative risk (RR) of an influenza infection occurring when the circulating strain does not match strains contained in the LAIV.

Group discussion

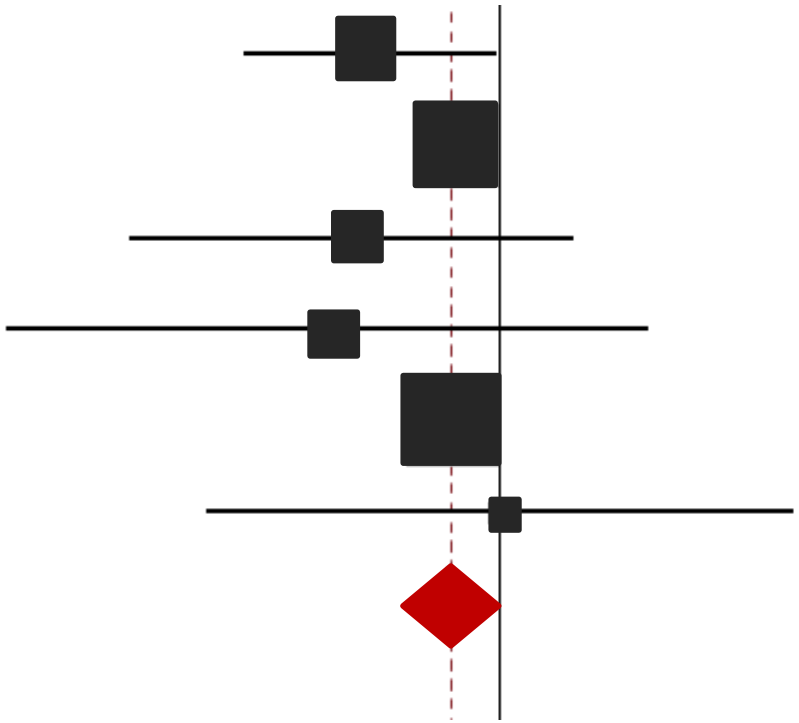
In the following example, which forest plot corresponds to the random-effects model?

Which forest plot corresponds to the random-effects model?

Forest plot 1



Forest plot 2



St. Michael's
Inspired Care.
Inspiring Science.

Exploring heterogeneity

- Characteristics of studies may be associated with the size of treatment effect

- For example,
  - risk of bias (e.g., adequate allocation concealment)
  - age group of patients
  - setting of study
  - dose of drug

- For discrete characteristics, can use subgroup analyses

- For discrete or continuous characteristics, can use meta-regression

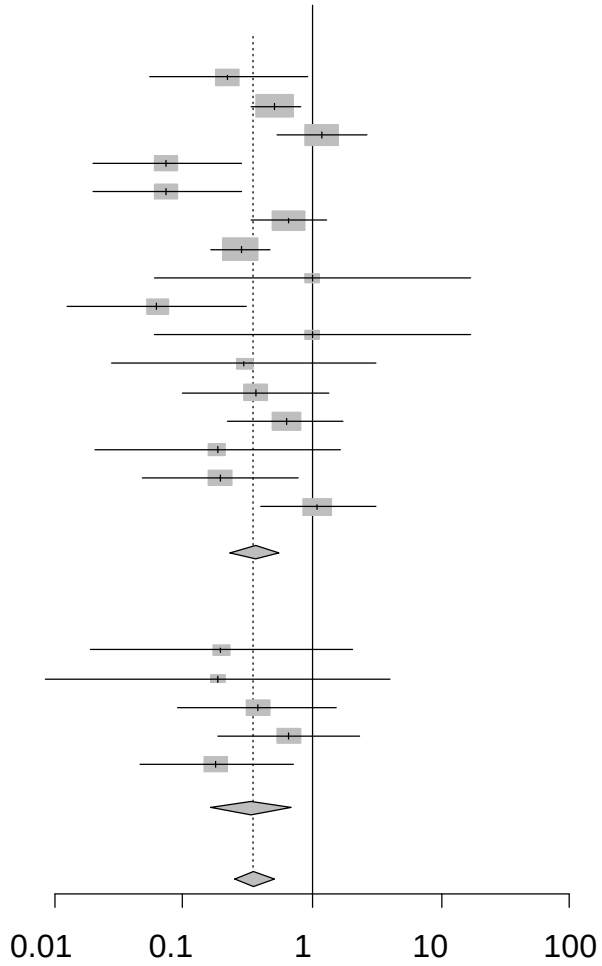
Group discussion

In the following example, do the results across the subgroup analyses differ?

Do the results across the subgroup analyses differ?

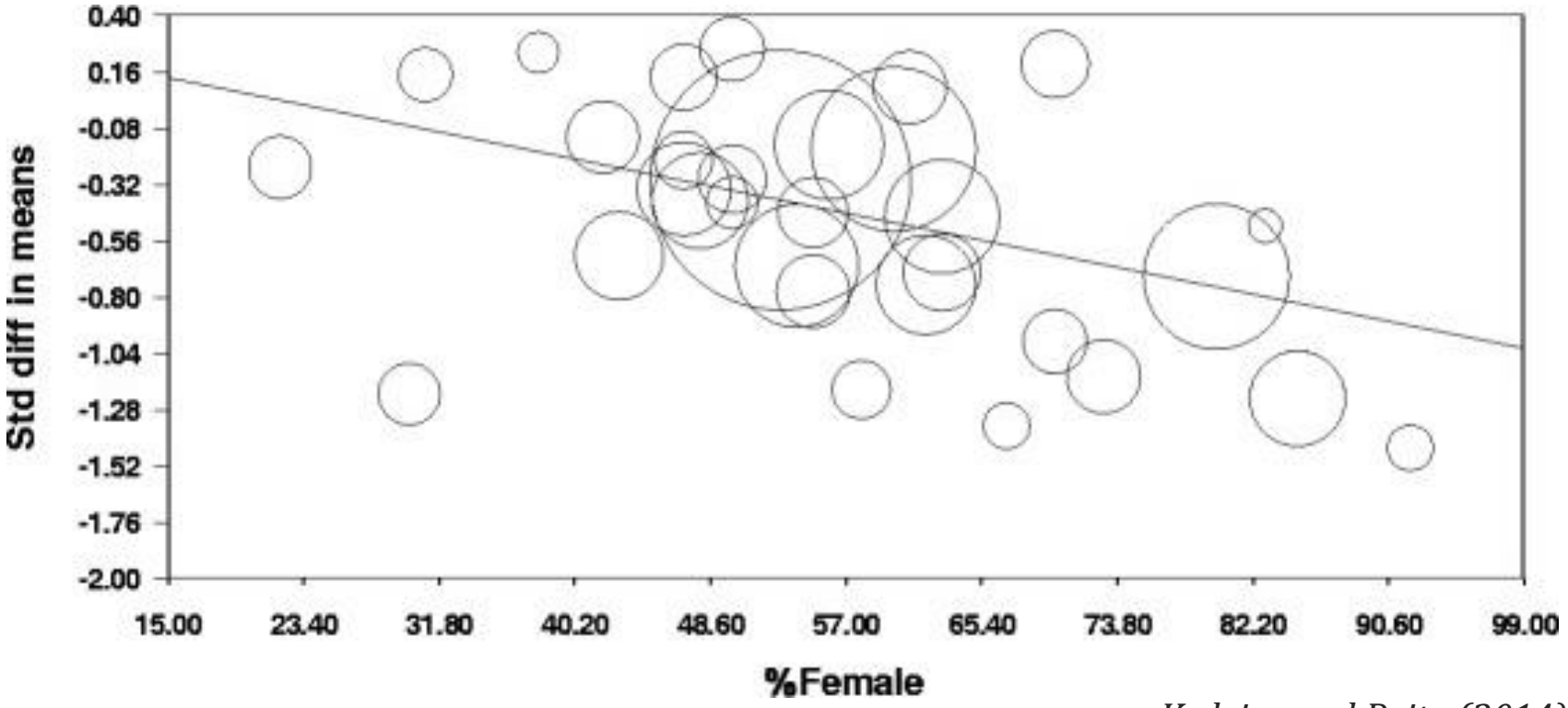
☐ YES ☐ NO

Study	Experimental Events	Experimental Total	Control Events	Control Total	OR	95%-CI	Weight
Subgroup 1							
119718	3	30	10	30	0.22	[0.05; 0.91]	4.4%
119782	58	165	85	165	0.51	[0.33; 0.79]	10.3%
119985	16	88	14	88	1.17	[0.53; 2.58]	7.8%
124097	19	40	37	40	0.07	[0.02; 0.28]	4.7%
124098	19	40	37	40	0.07	[0.02; 0.28]	4.7%
126278	40	70	47	70	0.65	[0.33; 1.30]	8.5%
126721	67	134	104	133	0.28	[0.16; 0.48]	9.7%
R59	1	33	1	33	1.00	[0.06; 16.69]	1.5%
120183	2	30	16	30	0.06	[0.01; 0.31]	3.7%
120185	1	30	1	30	1.00	[0.06; 16.76]	1.5%
121278	1	20	3	20	0.30	[0.03; 3.15]	2.0%
123155	4	30	9	30	0.36	[0.10; 1.33]	4.8%
123482	8	36	13	41	0.62	[0.22; 1.71]	6.3%
124081	1	50	5	50	0.18	[0.02; 1.63]	2.3%
124213	3	30	11	30	0.19	[0.05; 0.78]	4.4%
S2	9	42	9	45	1.09	[0.39; 3.08]	6.2%
Random effects model	**868**			**875**	**0.36**	**[0.23; 0.55]**	**82.8%**
Subgroup 2							
124295	1	15	4	15	0.20	[0.02; 2.02]	2.1%
125714	0	25	2	25	0.18	[0.01; 4.04]	1.3%
121423	4	20	8	20	0.38	[0.09; 1.54]	4.4%
122761	5	30	7	30	0.66	[0.18; 2.36]	4.9%
123056	5	20	13	20	0.18	[0.05; 0.70]	4.6%
Random effects model	**110**			**110**	**0.33**	**[0.16; 0.68]**	**17.2%**
Random effects model	**978**			**985**	**0.35**	**[0.24; 0.51]**	**100.0%**



Meta-regression - example

The treatment (high-frequency left repetitive transcranial magnetic stimulation) has higher antidepressant effect (against sham) in studies with higher proportion of female patients for the treatment of major depression



Kedzior and Reitz (2014)

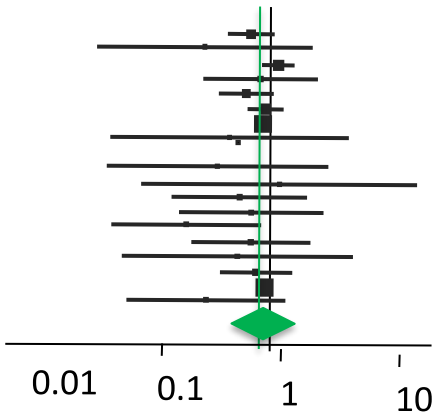
St. Michael's
Inspired Care.
Inspiring Science.

Interpreting the meta-analysis result

- Conventional Interpretations
  1. Statistical (and Clinical) Significance and Direction
  2. Magnitude of the pooled estimate
  3. Width of the confidence interval

- Heterogeneity
  - Excessive heterogeneity challenges the meaning of the meta-analysis result

- Quality of the included studies



St. Michael's
Inspired Care.
Inspiring Science.

Resources

- *Cochrane Handbook for Systematic Reviews of Interventions*
  - Higgins and Green (eds); Wiley 2008, updated online
- *Introduction to Meta-analysis*
  - Borenstein, Hedge, Higgins and Rothwell; Wiley 2009
- *Meta-Analysis of Controlled Clinical Trials*
  - Whitehead; Wiley 2002
- *Handbook of Research Synthesis and Meta-analysis*
  - Cooper, Hedges and Valentine; Sage 2009
- *Methods for Meta-Analysis on Medical Research*
  - Alex J.Sutton et al.,John Wiley & Sons, Ltd. (2000)
- *Statistical Methods for Meta-Analysis*
  - Larry V. Hedges and Ingrim Olgin, Academic Press, Inc. (1985)

References:

- Bowden J, Tierney JF, Copas AJ, Burdett S. 2011. Quantifying, displaying and accounting for heterogeneity in the meta-analysis of RCTs using standard and generalised Q statistics. BMC Medical Research Methodology 11: 41.
- Borrenstein M, Higgins JPT, Hedges LV, Rothstein H. Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. Res. Syn. Meth., 2017 (early view)
- DerSimonian R, Laird NM. Meta-analysis in clinical trials. Controlled Clinical Trials 1986. 7: 177–188.
- Higgins JPT, Thompson SG. 2002. Quantifying heterogeneity in a meta-analysis. Statistics in Medicine 21: 1539–1558.
- Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003. 327(7414):557-60.
- Higgins JPT, Thompson SG. Controlling the risk of spurious results from meta-regression. Stat Med 2004. 23(11):1663-82.

Resources

- Kedzior KK, Reitz SK. Short-term efficacy of repetitive transcranial magnetic stimulation (rTMS) in depression-reanalysis of data from meta-analyses up to 2010. BMC Psychol. 2014;2(1):39
- Riley RD, Higgins JPT, Deeks JJ. Interpretation of random effects meta-analyses. BMJ. 2011. 342: d549.
- Thompson SG. Controversies in meta-analysis: the case of the trials of serum cholesterol reduction. Stat Methods Med Res 1993; 2: 173-192
- Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. Br Med J 1994; 309: 1351-1355
- Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002; 21: 1559-1574
- Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Stat Med 1999; 18: 2693-2708
- Tricco AC, Chit A, Soobiah C, Hallett D, Meier G, Chen MH, Tashkandi M, Bauch CT, Loeb M. Comparing influenza vaccine efficacy against mismatched and matched strains: a systematic review and meta-analysis. BMC Med. 2013; 11:153.
- Tricco AC, Antony J, Ivers NM, Ashoor HM, Khan PA, Blondal E, Ghassemi M, MacDonald H, Chen MH, Ezer LK, Straus SE. Effectiveness of quality improvement strategies for coordination of care to reduce use of health care services: a systematic review and meta-analysis. CMAJ. 2014. 186(15):E568-78.
- Tricco AC, Soobiah C, Blondal E, Veroniki AA, Khan PA, Vafaei A, Ivory J, Strifler L, Ashoor H, MacDonald H, Reynen E, Robson R, Ho J, Ng C, Antony J, Mrklas K, Hutton B, Hemmelgarn BR, Moher D, Straus SE. Comparative safety of serotonin (5-HT3) receptor antagonists in patients undergoing surgery: a systematic review and network meta-analysis. BMC Med. 2015;13:142.
- Veroniki, A. A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., Kuss, O., Higgins, J. PT., Langan, D., and Salanti, G. Methods to estimate the between-study variance and its uncertainty in meta-analysis. Res. Syn. Meth., 2016 7: 55–79.
- Veroniki AA., Huedo-Medina T., Fountoulakis K Appraising between-study homogeneity, small study effects, moderators, and confounders.Umbrella reviews. Biondi-Zoccai G (ed). 2016.

Acknowledgements

- Dr. Georgia Salanti
- Prof. Julian P.T. Higgins
- Dr. Andrea C. Tricco
- Ms. Huda Ashoor
- Ms. Jesmin Antony
- Ms. Melissa Courvoisier

St. Michael's
Inspired Care.
Inspiring Science.

Questions?

St. Michael's
Inspired Care.
Inspiring Science.

Thank you for your attention!

Dr. Areti Angeliki Veroniki

Scientist: Knowledge Translation Program, Li Ka Shing Knowledge Institute of St. Michael's Hospital

VeronikiA@smh.ca

St. Michael's
Inspired Care.
Inspiring Science.