23 October 2024

Cochrane Methods

Introduction to new random-effects methods in RevMan

Areti-Angeliki Veroniki, MSc, PhD
St. Michael's Hospital, Unity Health Toronto
IHPME, University of Toronto

Jo McKenzie, MSc, PhD
Methods for Evidence Synthesis Unit,
School of Public Health and Preventive Medicine, Monash University

MONASH University

Institute of Health Policy, Management and Evaluation
UNIVERSITY OF TORONTO

KNOWLEDGE TRANSLATION PROGRAM

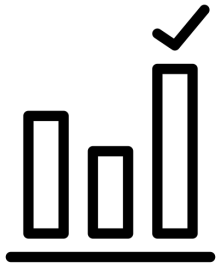
UNITY HEALTH TORONTO

Objective

To introduce the new random-effects methods being implemented in RevMan

Outline

- Process used to develop the recommendations for the random-effects meta-analysis methods
- Recap the random-effects model
- Outline the new random-effects methods, recommendations for when to use them (and why), and how they may impact the results (via example)
  - Heterogeneity estimator (and confidence interval method)
  - Confidence interval method for the summary mean effect
  - Prediction interval
- What to write in a protocol
- Questions



Created by Berkah Icon from Noun Project

Process used to develop and implement recommendations

…

2022 - 2020

1. Updated systematic reviews of statistical simulation studies examining the performance of heterogeneity estimators and CI methods for summary effect size [Veroniki 2016, Veroniki 2019]
2. Examined the impact of adopting different methods when applied to meta-analyses in the Cochrane Library
3. Convened multiple meetings to review and discuss the evidence and form recommendations

2024 - 2022

4. Submitted recommendations and evidence to the Methods Executive (Jan 2022)
   Cochrane endorsed the recommendations (Sep 2022)
   Revision of recommendations (May 2024)

Team developing recommendations:

- Areti-Angeliki Veroniki | Unity Health Toronto, University of Toronto
- Dean Langan | University College London
- Simon Turner | Monash University
- Mark Simmonds | University of York
- Anna Chaimani | Université Paris Cité
- Kerry Dwan | formally Cochrane
- Joanne McKenzie | Monash University

Experience:

- Co-convenors of the Cochrane Statistical Methods Group
- Led systematic reviews of statistical simulation studies, and undertaken simulation studies, examining random-effects methods
- Led empirical evaluations examining the impact of using different methods
- Cochrane Methods Support Unit Lead and Statistical Editor

Process used to develop and implement recommendations

…

2022 - 2020

1. Updated systematic reviews of statistical simulation studies examining the performance of heterogeneity estimators and CI methods for summary effect size [Veroniki 2016, Veroniki 2019]
2. Examined the impact of adopting different methods when applied to meta-analyses in the Cochrane Library
3. Convened multiple meetings to review and discuss the evidence and form recommendations

2024 - 2022

4. Submitted recommendations and evidence to the Methods Executive (Jan 2022)
   Cochrane endorsed the recommendations (Sep 2022)
   Revision of recommendations (May 2024)

Current - 2023

5. Integration of the methods into RevMan
6. Implementation activities (e.g. updating Chapter 10 Cochrane Handbook, webinars, training materials, workshops)

Cochrane Methods Implementation Editor:
- Ingrid Arévalo-Rodriguez

Cochrane IT development and Infrastructures:
- Rebecka Hall | RevMan Product Owner
- Gert van Valkenhoef | Head
- Rasmus Moustgaard | Senior Systems Architect
- + Others

Specialist statistical advice from:
- Julian Higgins | University of Bristol
- Wolfgang Viechtbauer | metafor package creator

Cochrane Statistical Methods Group links:
- Areti-Angeliki Veroniki | Unity Health Toronto, University of Toronto
- Joanne McKenzie | Monash University

Testing:
- *Simon Turner* | Monash University

Process used to develop and implement recommendations

…

2022 - 2020

1. Updated systematic reviews of statistical simulation studies examining the performance of heterogeneity estimators and CI methods for summary effect size [Veroniki 2016, Veroniki 2019]
2. Examined the impact of adopting different methods when applied to meta-analyses in the Cochrane Library
3. Convened multiple meetings to review and discuss the evidence and form recommendations

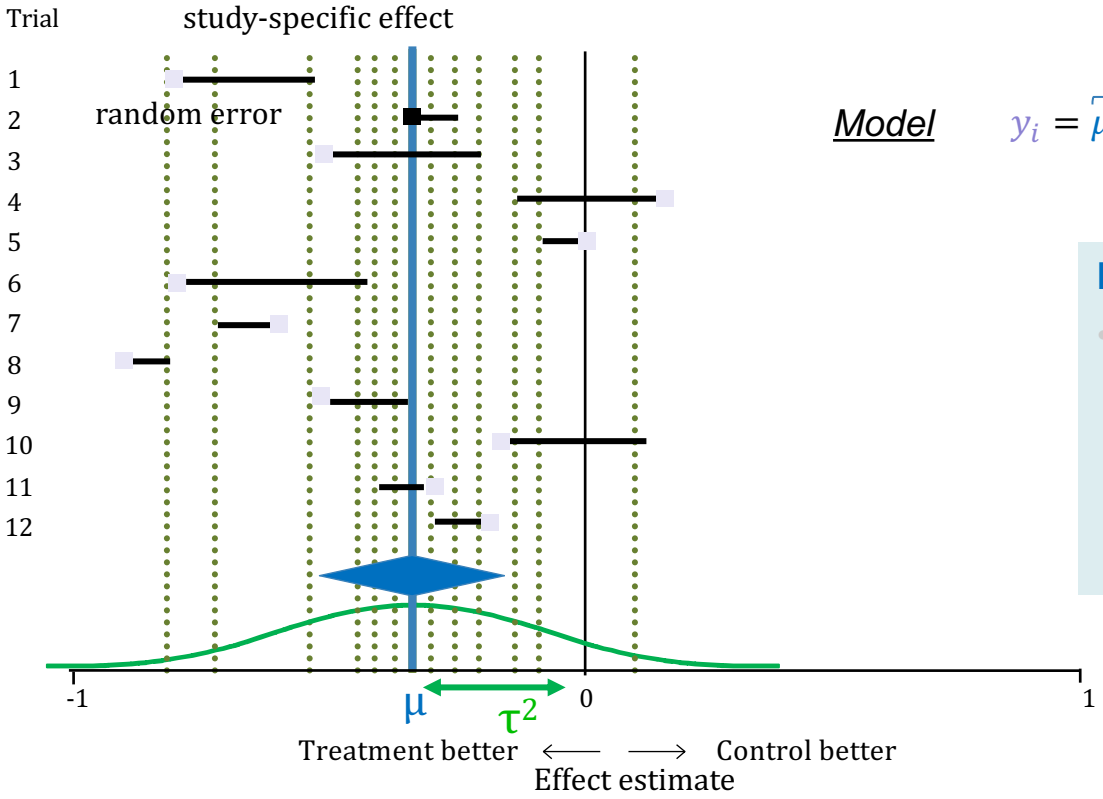
2024 - 2022

4. Submitted recommendations and evidence to the Methods Executive (Jan 2022)
   Cochrane endorsed the recommendations (Sep 2022)
   Revision of recommendations (May 2024)

Current - 2023

5. Integration of the methods into RevMan
6. Implementation activities (e.g. updating Chapter 10 Cochrane Handbook, webinars, training materials, workshops)

Random-effects meta-analysis



Model $y_i = \overbrace{\mu + u_i}^{\theta_i} + \varepsilon_i$

$u_i \sim N(0, \tau^2)$

$\varepsilon_i \sim N(0, s_i^2)$

Main Assumption:

- The observed study-specific effects estimate different *true* effects, which are related and come from the same distribution

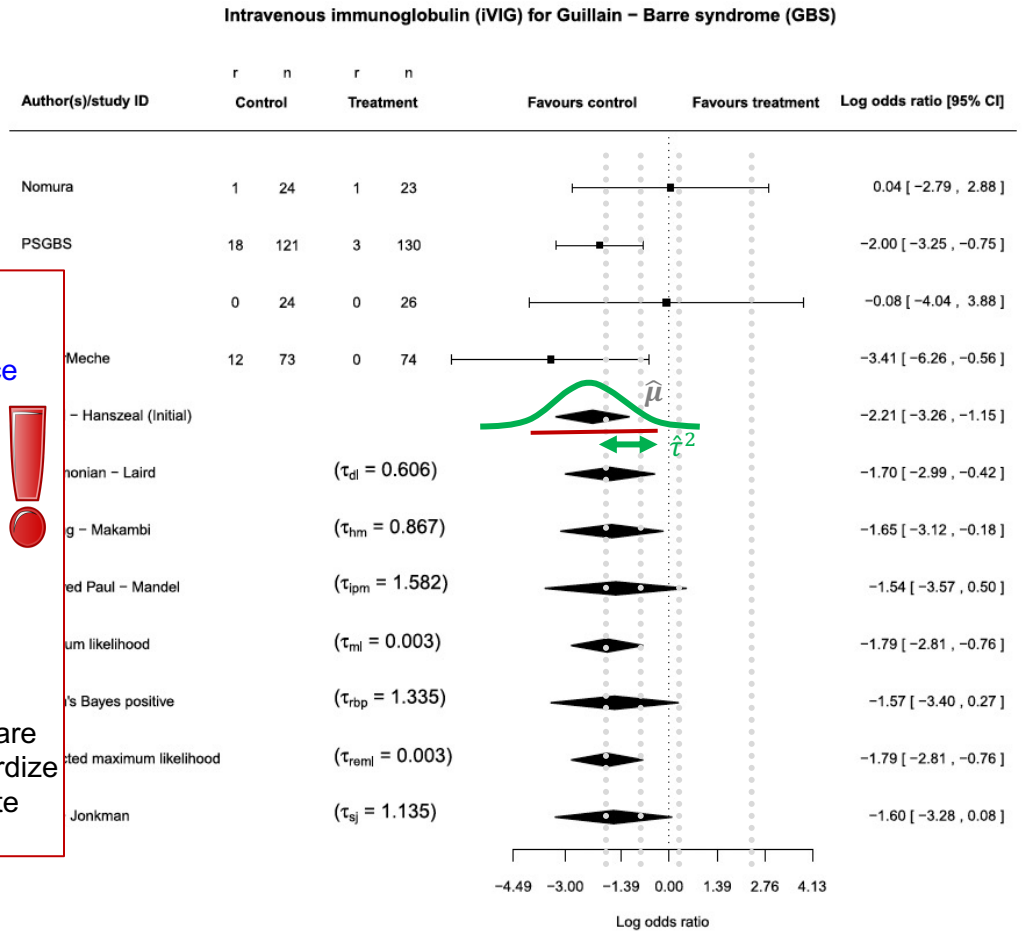
Random-effects meta-analysis model

The choice of the method for estimating

- between-study variance (heterogeneity) and its uncertainty
- uncertainty for the summary effect size

is important when conducting a meta-analysis

When inappropriate methods are used, this can seriously jeopardize results, leading to inappropriate conclusions



Under the random-effects model, we can estimate a number of parameters and calculate several statistics, including:

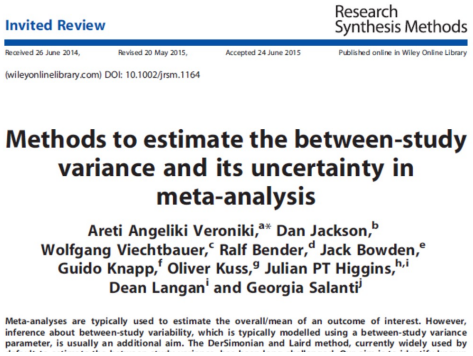
- Average (summary) effect ($\hat{\mu}$), along with a CI
- Between-study variance ($\hat{\tau}^2$), along with a CI
- **Prediction interval** (predicted range for the true treatment effect in an individual study)
- + others (e.g., I^2, H^2)

Random-effects meta-analysis model

- **DerSimonian & Laird (DL)** is the frequently random-effects meta-analysis method used
- DL is a method of moments estimator of τ^2
- The Wald-type normal distribution is used to calculate a CI for the summary effect
- DL with the Wald-type normal distribution is the **only** random-effects method implemented in **RevMan**
- Different estimators of heterogeneity (τ^2) and methods to calculate uncertainty in the summary effect exist
- For any particular meta-analysis, the estimated parameters (e.g. summary effect, heterogeneity variance) may differ depending on the method used

Which is the most appropriate method to use?

Work conducted on behalf of the Cochrane Statistical Methods Group



Invited Review

Research Synthesis Methods

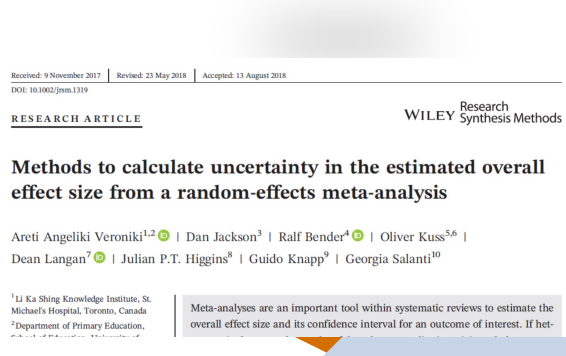
Received 26 June 2014, Revised 20 May 2015, Accepted 24 June 2015 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jrsm.1164

Methods to estimate the between-study variance and its uncertainty in meta-analysis

Areti Angeliki Veroniki,[a*] Dan Jackson,[b] Wolfgang Viechtbauer,[c] Ralf Bender,[d] Jack Bowden,[e] Guido Knapp,[f] Oliver Kuss,[g] Julian PT Higgins,[h,i] Dean Langan[i] and Georgia Salanti[j]

Meta-analyses are typically used to estimate the overall/mean of an outcome of interest. However, inference about between-study variability, which is typically modelled using a between-study variance parameter, is usually an additional aim. The DerSimonian and Laird method, currently widely used by



Received: 9 November 2017 | Revised: 23 May 2018 | Accepted: 13 August 2018

DOI: 10.1002/jrsm.1319

RESEARCH ARTICLE

WILEY Research Synthesis Methods

Methods to calculate uncertainty in the estimated overall effect size from a random-effects meta-analysis

Areti Angeliki Veroniki[1,2] | Dan Jackson[3] | Ralf Bender[4] | Oliver Kuss[5,6] | Dean Langan[7] | Julian P.T. Higgins[8] | Guido Knapp[9] | Georgia Salanti[10]

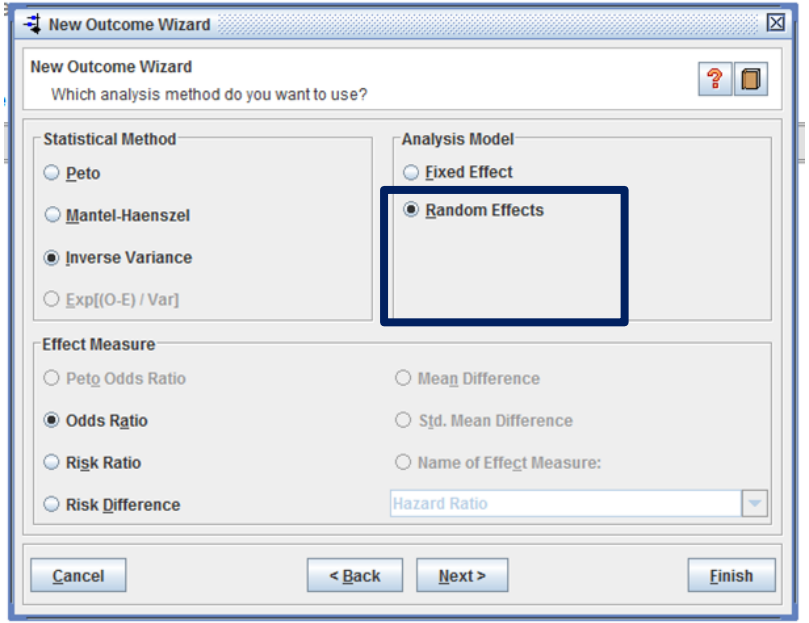
[1]Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Canada

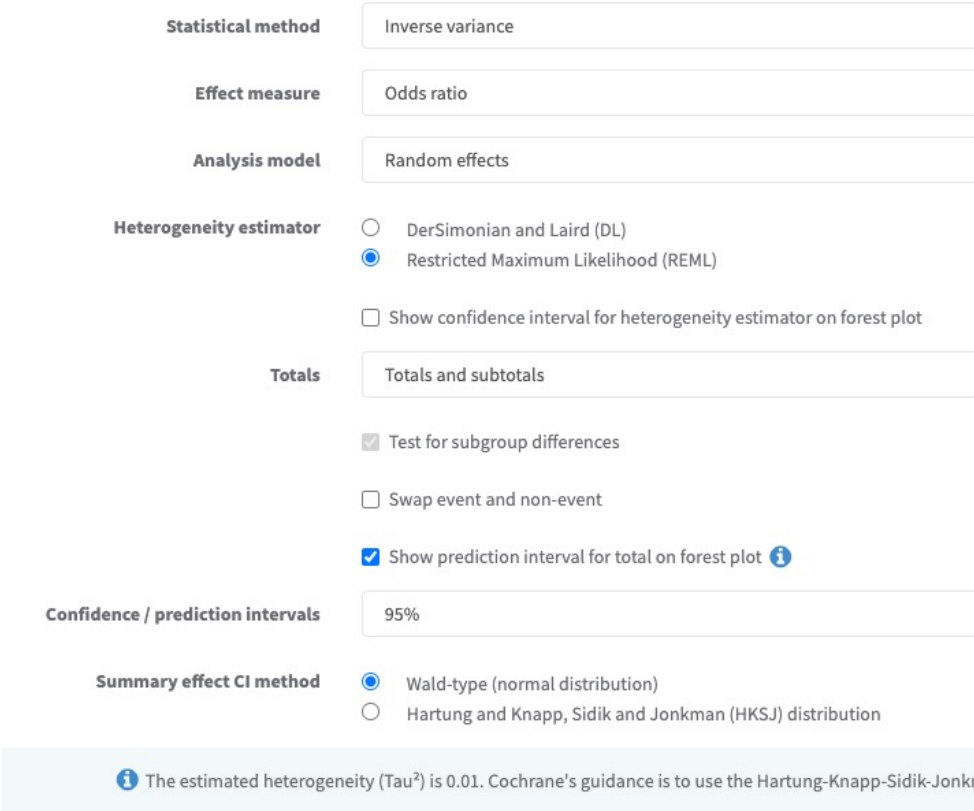
[2]Department of Primary Education,

Meta-analyses are an important tool within systematic reviews to estimate the overall effect size and its confidence interval for an outcome of interest. If het-

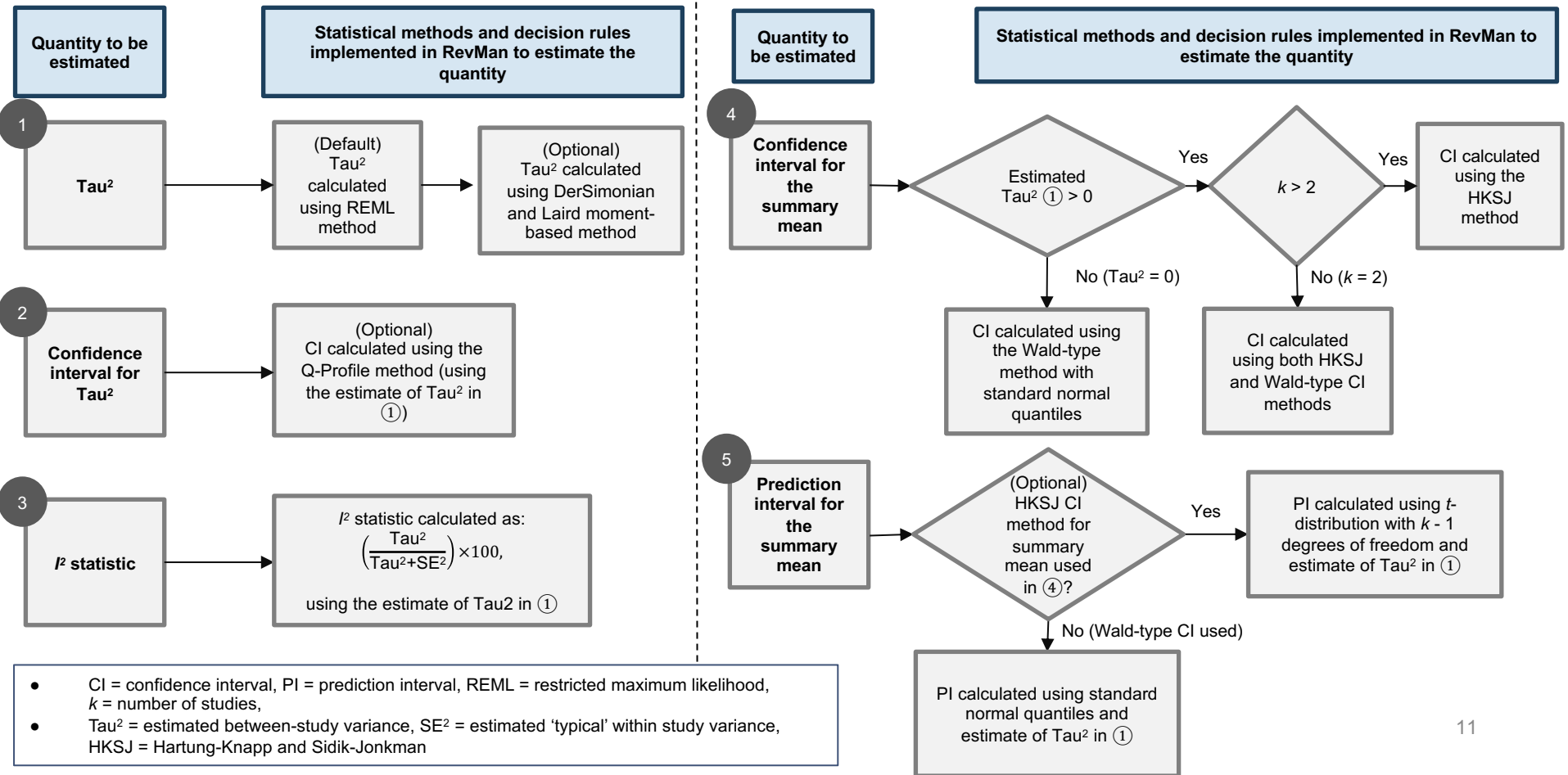
Updating RevMan

RevMan



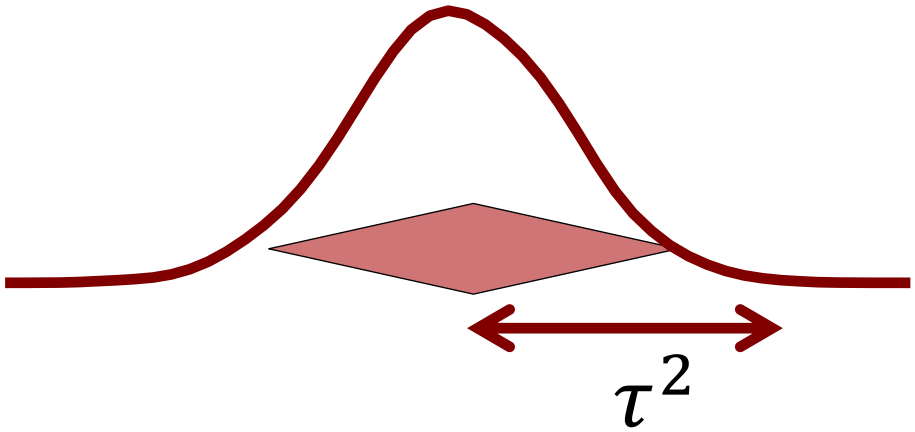


Random-effects methods implemented in RevMan



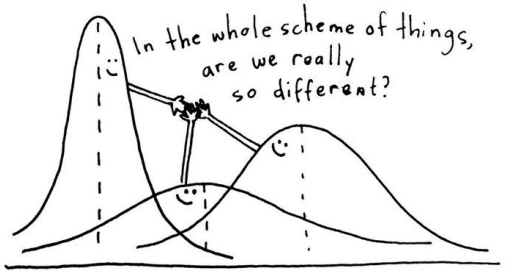
- CI = confidence interval, PI = prediction interval, REML = restricted maximum likelihood, k = number of studies,
- Tau^2 = estimated between-study variance, SE^2 = estimated 'typical' within study variance, HKSJ = Hartung-Knapp and Sidik-Jonkman

1. Inference on the heterogeneity



Inference on the heterogeneity

Quantity to be estimated	Statistical methods and decision rules implemented in RevMan to estimate the quantity	
1 Tau²	(**Default**) Tau² calculated using REML method	(**Optional**) Tau² calculated using DerSimonian and Laird moment-based method
2 Confidence interval for Tau²	(**Optional**) CI calculated using the Q-Profile method (using the estimate of Tau² in ①)	
3 I^2 statistic	I^2 statistic calculated as: $\left(\frac{\text{Tau}^2}{\text{Tau}^2+\text{SE}^2}\right)\times 100$, using the estimate of Tau2 in ①	



Recommendations based on published studies

An empirical study using 57,397 Cochrane meta-analyses with $k \geq 2$ showed that:
→ The mean τ^2 is higher than generally assumed but fails to be detected, especially for small k!

Kontopantelis et al. 2013

A descriptive analysis of Cochrane systematic reviews found that 75% of meta-analyses contained 5 or fewer studies

Davey et al. 2011

The majority of the pairwise meta-analyses have:

$k \leq 10$

Turner et al 2012
Pullenayegum et al 2011
Rhodes et al 2014

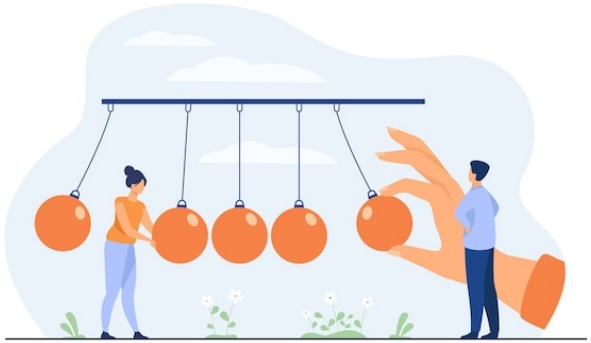
Problem for Cochrane reviews → few studies

- e.g. Langan 2015 median 4 [IQR 3-7]

Implications with different estimators for heterogeneity

According to simulation and empirical findings, the main factors (among others) that may affect the between-study variance estimation are:

- Number and size of studies included in the meta-analysis
- Magnitude of true heterogeneity
- Frequency of events (for dichotomous outcomes)

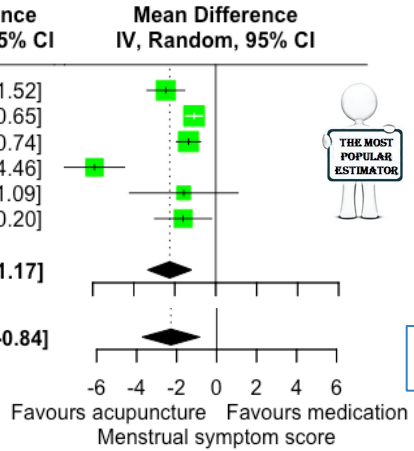


DL often underestimates heterogeneity

(particularly when the number of studies is small)

Acupuncture for dysmenorrhoea

Study	Intervention Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Ruan 2011	3.55	1.2100	30	5.99	2.2700	30	18.5%	-2.44 [-3.36; -1.52]
Peng 2012	1.77	0.9000	30	2.83	0.7000	30	20.8%	-1.06 [-1.47; -0.65]
Han 2012	1.70	1.1500	80	3.03	1.7100	40	20.1%	-1.33 [-1.92; -0.74]
Zhang 2013a	2.29	1.3300	30	8.21	3.8700	30	15.4%	-5.92 [-7.38; -4.46]
Qiao 2013	8.14	4.1600	60	9.70	5.5500	20	9.4%	-1.56 [-4.21; 1.09]
Wang 2014a	5.53	2.8000	30	7.12	2.7000	30	15.8%	-1.59 [-2.98; -0.20]
Random effects model (DL)			**260**			**180**	**100.0%**	**-2.25 [-3.33; -1.17]**
Heterogeneity: Tau^2 = 1.4274; Chi^2 = 43.81, df = 5 (P < 0.0001)								
Random effects model (REML)			**260**			**180**	**100.0%**	**-2.29 [-3.73; -0.84]**
Heterogeneity: Tau^2 = 2.8153; Chi^2 = 43.81, df = 5 (P < 0.0001)								



DL: $\hat{\tau}^2$= 1.427 ; I^2= 89%

REML: $\hat{\tau}^2$= 2.815; I^2= 94%

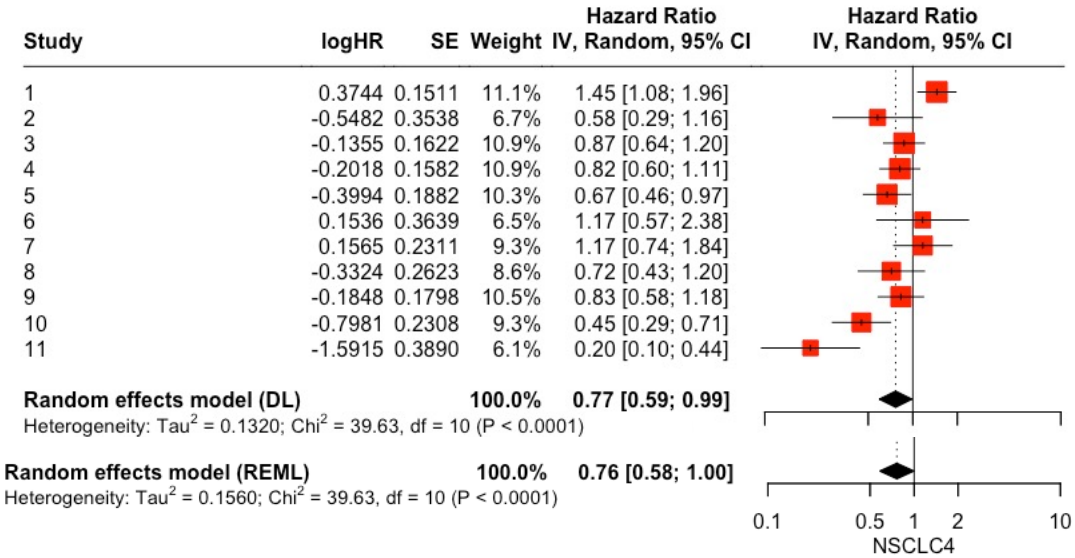
The amount of between-study variance can be estimated, but estimates are usually imprecise

Obtain a CI for τ^2!

95% CI for $\hat{\tau}^2$: [0.824, 19.515]

Smith et al CDSR 2016: https://doi.org/10.1002/14651858.CD006930.pub3

Simulations have shown that REML provides more accurate estimates with less bias



DL: $\hat{\tau}^2$= 0.132 ; I^2= 75%

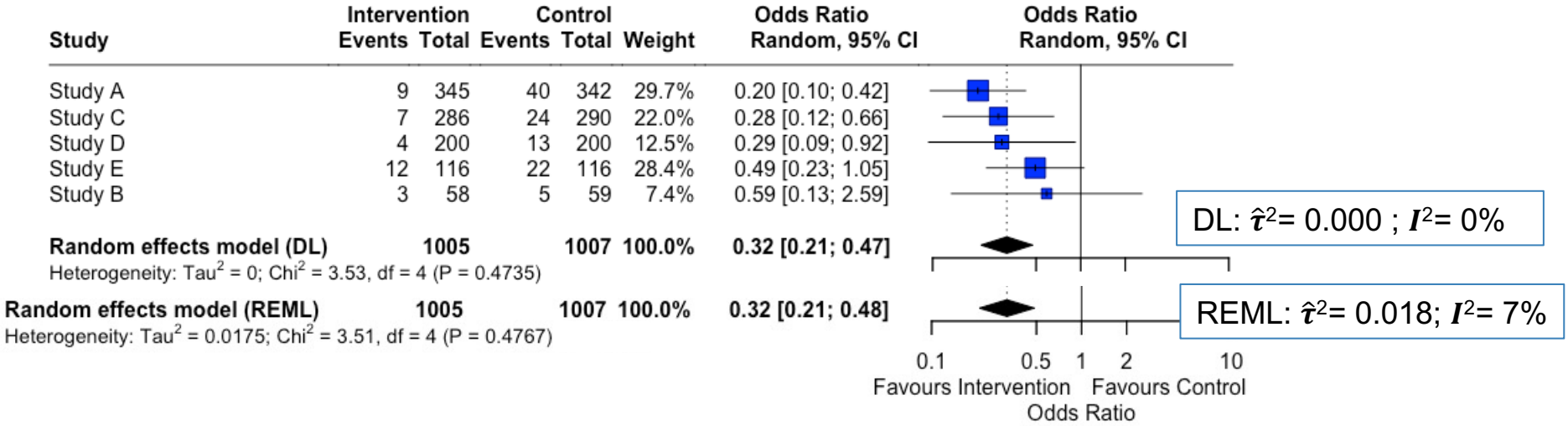
REML: $\hat{\tau}^2$= 0.156; I^2= 78%

Confidence interval for Tau²

95% CI for $\hat{\tau}^2$: [0.052 0.787]

Bowden et *al.* BMC Medical Research Methodology 11: 41, 2011. DOI:10.1186/1471-2288-11-41

DL frequently estimates tau=0



DL: $\hat{\tau}^2$= 0.000 ; I^2= 0%

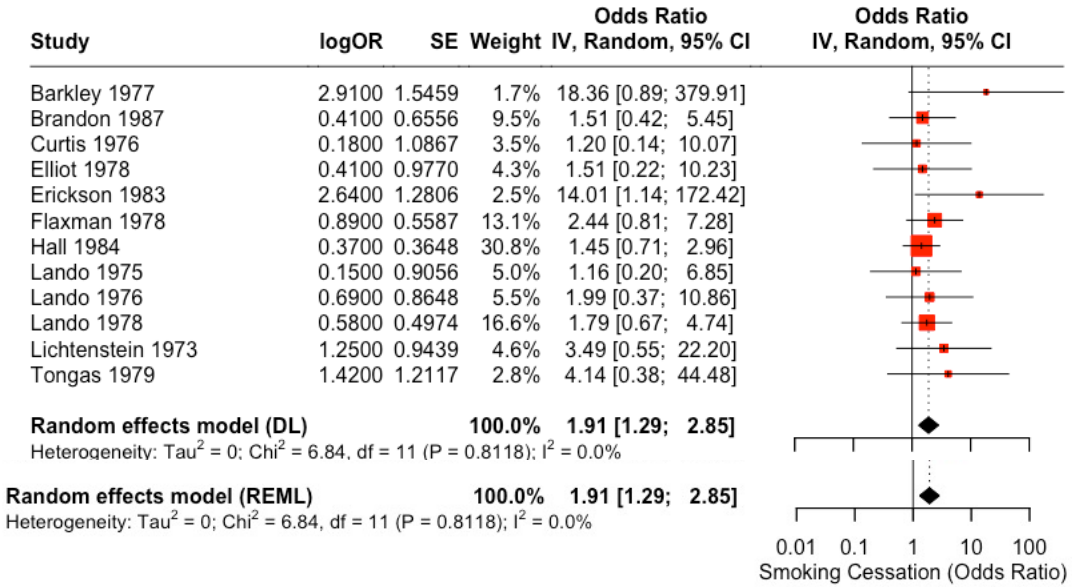
REML: $\hat{\tau}^2$= 0.018; I^2= 7%

Confidence interval for Tau²

95% CI for $\hat{\tau}^2$: [0.000 1.298]

When the number of studies increases DL tends to agree with REML

Aversive smoking for smoking cessation



Confidence interval for Tau²

95% CI for $\hat{\tau}^2$: [0.000 0.904]

DL: $\hat{\tau}^2$= 0.000 ; I^2= 0%

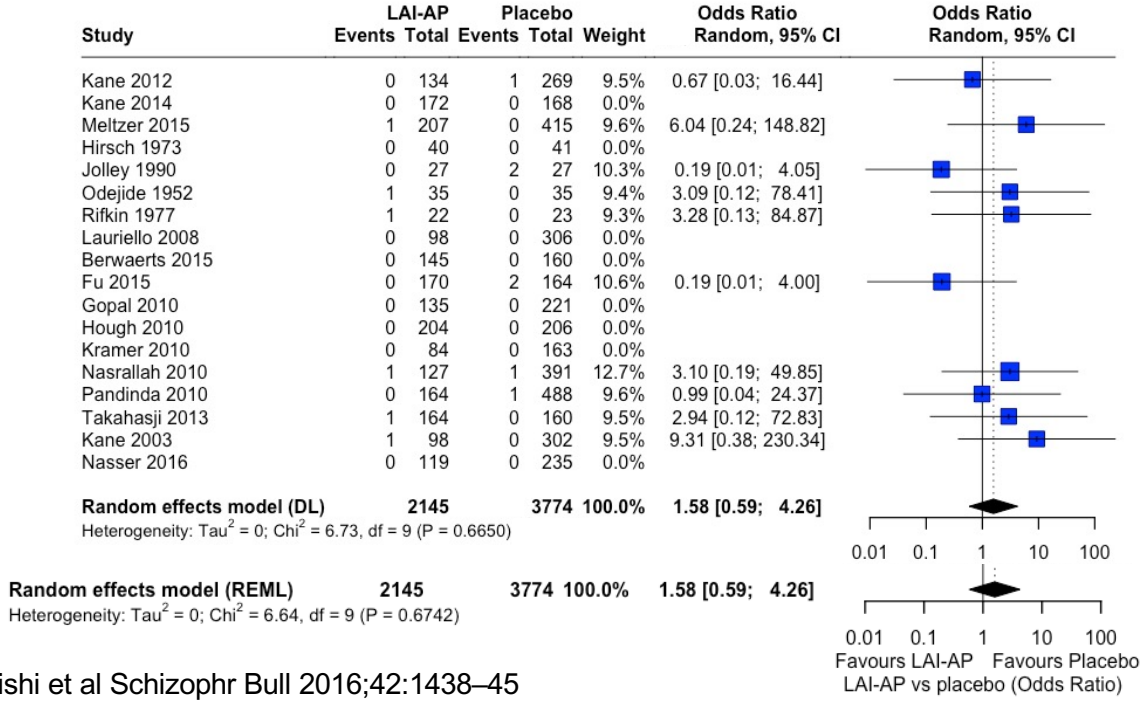
REML: $\hat{\tau}^2$= 0.000; I^2= 0%

Jackson and White, Biometrical 2018

In case of **rare events**, both DL and REML tend to underestimate heterogeneity

All-cause mortality in antipsychotics

Long-acting injectable antipsychotics (LAI-AP) vs Placebo

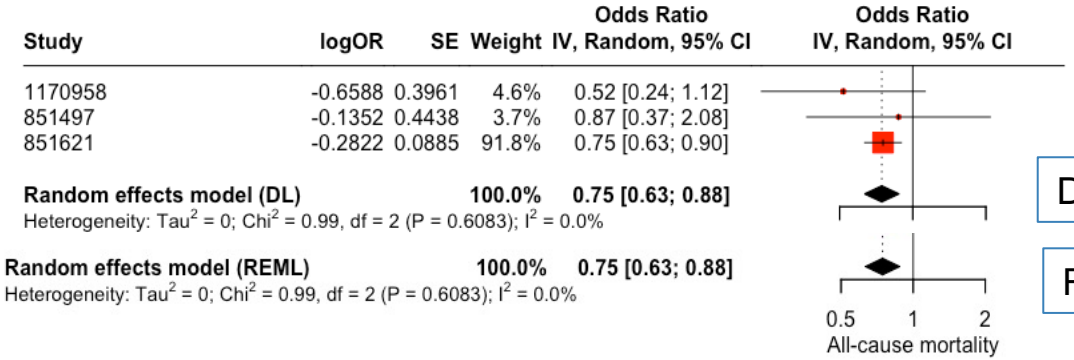


DL: $\hat{\tau}^2$= 0.000 ; I^2= 0%

REML: $\hat{\tau}^2$= 0.000; I^2= 0%

Kishi et al Schizophr Bull 2016;42:1438–45

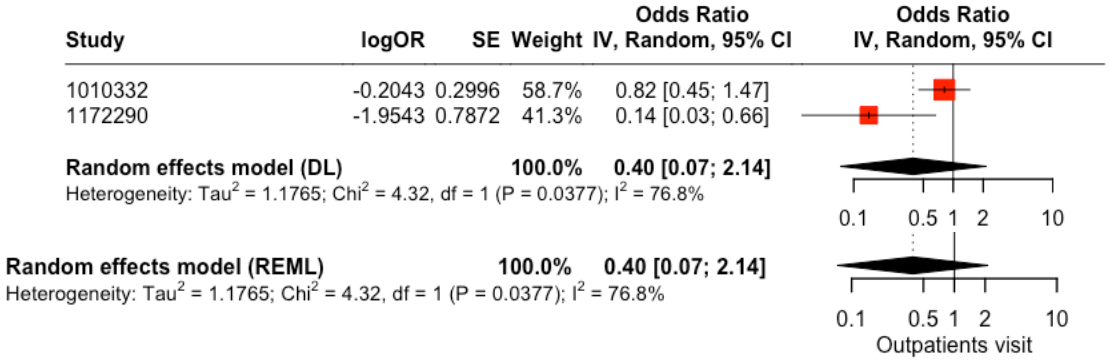
For very few studies both DL and REML tend to underestimate heterogeneity



DL: $\hat{\tau}^2$= 0.00 ; I^2= 0%

REML: $\hat{\tau}^2$= 0.00 ; I^2= 0%

95% CI for $\hat{\tau}^2$: [0.000 2.717]

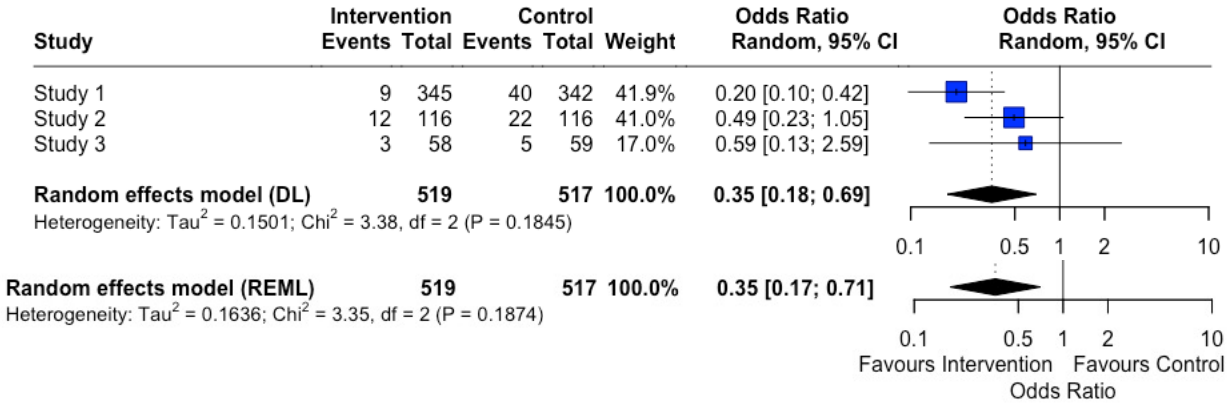


DL: $\hat{\tau}^2$= 1.176 ; I^2= 76.8%

REML: $\hat{\tau}^2$= 1.176 ; I^2= 76.8%

95% CI for $\hat{\tau}^2$: [0.000 >100]

For very few studies both DL and REML tend to underestimate heterogeneity but usually REML performs best



DL: $\hat{\tau}^2$= 0.151 ; I^2= 40%

REML: $\hat{\tau}^2$= 0.1636; I^2= 43%

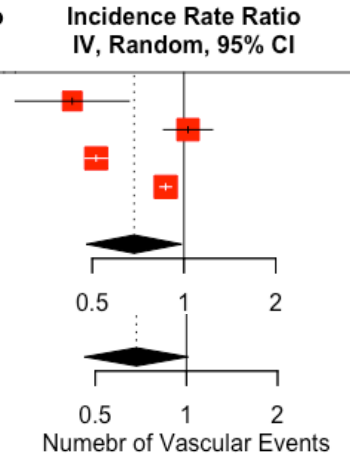
Confidence interval for Tau²

95% CI for $\hat{\tau}^2$: [0.000 12.707]

For very few studies both DL and REML tend to underestimate heterogeneity but usually REML performs best

(particularly when heterogeneity is high)

Study	logIRR	SE	Weight	Incidence Rate Ratio IV, Random, 95% CI
1171692	-0.8442	0.2193	19.8%	0.43 [0.28; 0.66]
1172124	0.0304	0.0945	25.7%	1.03 [0.86; 1.24]
851497	-0.6637	0.0435	27.1%	0.51 [0.47; 0.56]
851621	-0.1383	0.0219	27.4%	0.87 [0.83; 0.91]
Random effects model (DL)			**100.0%**	**0.69 [0.48; 0.98]**
Heterogeneity: $Tau^2 = 0.1223$; $Chi^2 = 132.36$, df = 3 ($P < 0.0001$)				
Random effects model (REML)			**100.0%**	**0.68 [0.46; 1.01]**
Heterogeneity: $Tau^2 = 0.1488$; $Chi^2 = 132.36$, df = 3 ($P < 0.0001$)				



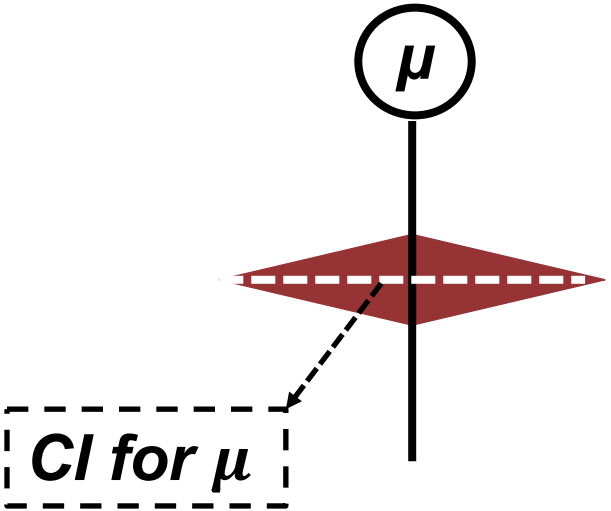
DL: $\hat{\tau}^2 = 0.122$; $I^2 = 97.7\%$

REML: $\hat{\tau}^2 = 0.149$; $I^2 = 98.1\%$

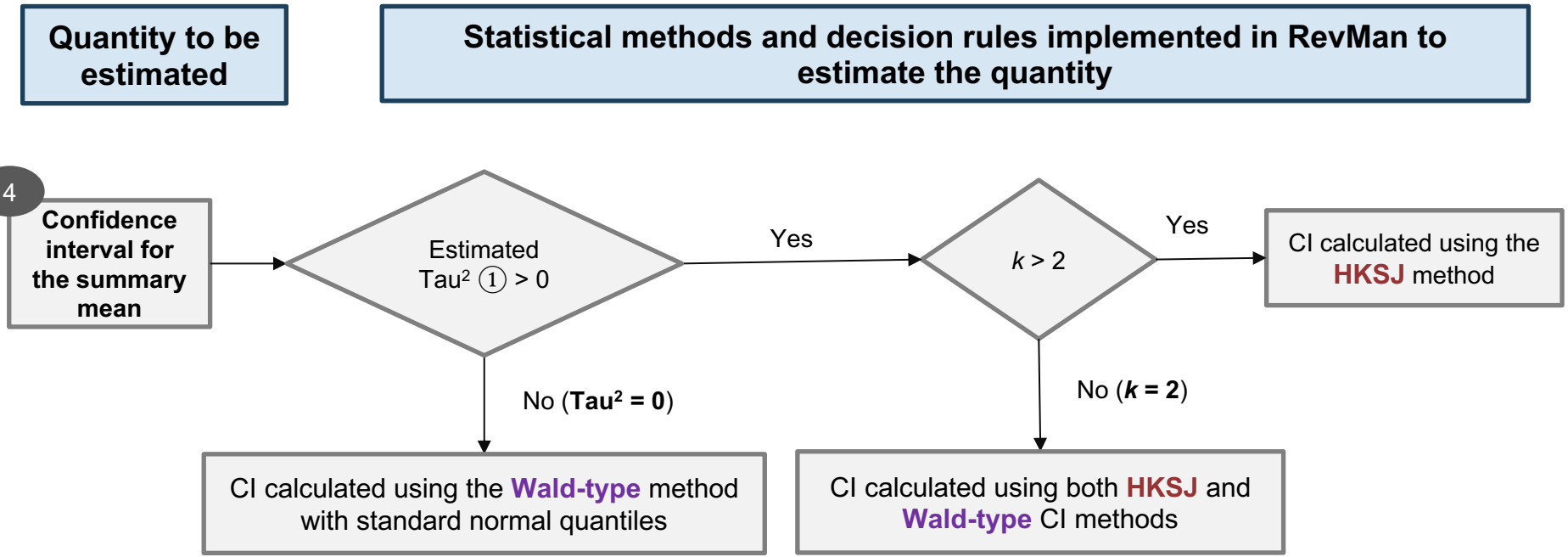
Confidence interval for Tau²

95% CI for $\hat{\tau}^2$: [0.039 2.391]

2. Inference on the summary mean effect



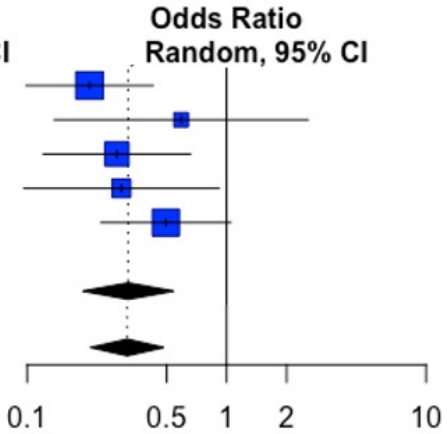
Inference on the summary mean effect

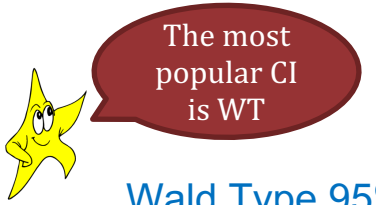


WT depends on the number of studies
(For few studies the CIs for μ are too narrow)

Study	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Odds Ratio Random, 95% CI
Study A	9	345	40	342	29.0%	0.20 [0.10; 0.42]
Study B	3	58	5	59	7.9%	0.59 [0.13; 2.59]
Study C	7	286	24	290	22.2%	0.28 [0.12; 0.66]
Study D	4	200	13	200	13.1%	0.29 [0.09; 0.92]
Study E	12	116	22	116	27.9%	0.49 [0.23; 1.05]
Random effects model (HKSJ)		**1005**		**1007**	**100.0%**	**0.32 [0.19; 0.54]**
Random effects model (Wald Type)		**1005**		**1007**	**100.0%**	**0.32 [0.21; 0.48]**

Heterogeneity: $Tau^2 = 0.0175$; $Chi^2 = 3.51$, df = 4 (P = 0.48); $I^2 = 0\%$





standard normal distribution

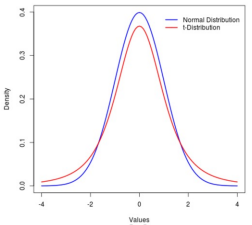
Wald Type 95% CI: $\hat{\mu} \pm 1.96\sqrt{var_{WT}(\hat{\mu})}$

[0.21, 0.48]

t-distribution

HKSJ 95% CI: $\hat{\mu} \pm t_{k-1,0.975}\sqrt{var_{HKSJ}(\hat{\mu})}$

[0.19, 0.54]

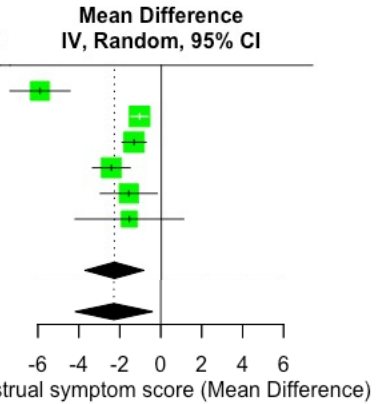


HKSJ on average produces wider CIs, but captures the true summary effect

Acupuncture for dysmenorrhoea

Study	Acupuncture Mean	SD	Total	Medication Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Zhang 2013a	2.29	1.3300	30	8.21	3.8700	30	16.1%	-5.92 [-7.38; -4.46]
Peng 2012	1.77	0.9000	30	2.83	0.7000	30	19.0%	-1.06 [-1.47; -0.65]
Han 2012	1.70	1.1500	80	3.03	1.7100	40	18.7%	-1.33 [-1.92; -0.74]
Ruan 2011	3.55	1.2100	30	5.99	2.2700	30	17.9%	-2.44 [-3.36; -1.52]
Wang 2014a	5.53	2.8000	30	7.12	2.7000	30	16.4%	-1.59 [-2.98; -0.20]
Qiao 2013	8.14	4.1600	60	9.70	5.5500	20	11.7%	-1.56 [-4.21; 1.09]
Random effects model (WT)			**260**			**180**	**100.0%**	**-2.29 [-3.73; -0.84]**
Random effects model (HKSJ)			**260**			**180**	**100.0%**	**-2.29 [-4.19; -0.39]**

Heterogeneity: $Tau^2 = 2.8153$; $Chi^2 = 43.81$, df = 5 ($P < 0.0001$); $I^2 = 88.6\%$



$\hat{\tau}^2 = 2.815$

WT: [-3.73, -0.84]

HKSJ: [-4.19, -0.39]

With $var_{HKSJ}(\hat{\mu}) > var_{WT}(\hat{\mu})$, when $\hat{\tau}^2 > 0$

Wald Type 95% CI: $\hat{\mu} \pm 1.96\sqrt{var_{WT}(\hat{\mu})}$

[-3.73, -0.84]

HKSJ 95% CI: $\hat{\mu} \pm t_{k-1,0.975}\sqrt{var_{HKSJ}(\hat{\mu})}$

[-4.19, -0.39]

Smith et al CDSR 2016: https://doi.org/10.1002/14651858.CD006930.pub3

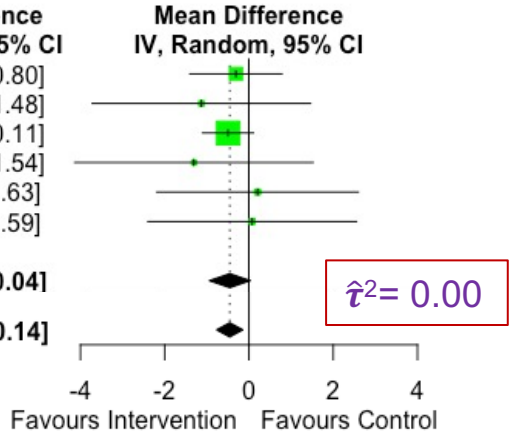
In the absence of heterogeneity: HKSJ < WT

Alcohol-related problems: up to 3 months

Social norms (SN) vs control

Study	Intervention Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Lewis 2008	2.33	3.8400	97	2.64	3.8900	90	20.0%	-0.31 [-1.42; 0.80]
Walters 2000	4.86	3.4800	14	6.00	3.1900	11	3.6%	-1.14 [-3.76; 1.48]
Werch 2000	2.20	3.1000	255	2.70	4.0000	266	65.4%	-0.50 [-1.11; 0.11]
Ju√°rez 2006	4.28	4.2100	21	5.60	5.0800	20	3.0%	-1.32 [-4.18; 1.54]
Geisner 2007	5.24	7.8900	88	5.03	8.5300	89	4.2%	0.21 [-2.21; 2.63]
Collins 2002	7.91	5.6900	47	7.83	6.6700	47	3.9%	0.08 [-2.43; 2.59]
Random effects model (Wald type)			**522**			**523**	**100.0%**	**-0.46 [-0.95; 0.04]**
Random effects model (HKSJ)			**522**			**523**	**100.0%**	**-0.46 [-0.77; -0.14]**

Heterogeneity: $Tau^2 = 0$; $Chi^2 = 1.16$, df = 5 (P = 0.95); $I^2 = 0\%$



Wald Type 95% CI: $\hat{\mu} \pm 1.96\sqrt{var_{WT}(\hat{\mu})}$

$[-0.46 \pm 1.96^* \sqrt{0.06}]$

[-0.95, 0.04]

HKSJ 95% CI: $\hat{\mu} \pm t_{k-1,0.975}\sqrt{var_{HKSJ}(\hat{\mu})}$

$[-0.46 \pm 2.57^* \sqrt{0.015}]$

[-0.77, -0.14]

Foxcroft et al CDSR 2015: https://pubmed.ncbi.nlm.nih.gov/25622306/

In the absence of heterogeneity: HKSJ < WT

(Irrespective of the number of studies)

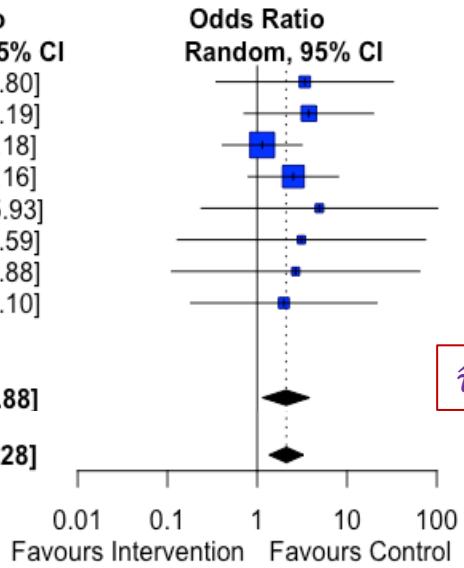
Respiratory distress syndrome

Sotiriadis et al CDSR 2018: https://pubmed.ncbi.nlm.nih.gov/30075059/

Antenatal corticosteroids vs no steroids

Study	Intervention Events	Intervention Total	Control Events	Control Total	Weight	Odds Ratio Random, 95% CI
Stutchfield 2005	3	66	1	73	7.1%	3.43 [0.35; 33.80]
Ahmed 2015	6	36	2	40	13.2%	3.80 [0.72; 20.19]
Nooh 2018	8	301	7	299	35.0%	1.14 [0.41; 3.18]
Nada 2016	10	611	4	616	27.2%	2.55 [0.79; 8.16]
Ahmed 2015	2	76	0	74	4.0%	5.00 [0.24; 105.93]
Stutchfield 2005	1	210	0	219	3.6%	3.14 [0.13; 77.59]
Stutchfield 2005	1	195	0	175	3.6%	2.71 [0.11; 66.88]
Nooh 2018	2	301	1	299	6.4%	1.99 [0.18; 22.10]
Ahmed 2015	0	112	0	114	0.0%	
Random effects model (Wald Type)		**1908**		**1909**	**100.0%**	**2.11 [1.15; 3.88]**
Random effects model (HKSJ)		**1908**		**1909**	**100.0%**	**2.11 [1.36; 3.28]**

Heterogeneity: $Tau^2 = 0$; $Chi^2 = 2.52$, df = 7 (P = 0.93); $I^2 = 0\%$



$\hat{\tau}^2 = 0.00$

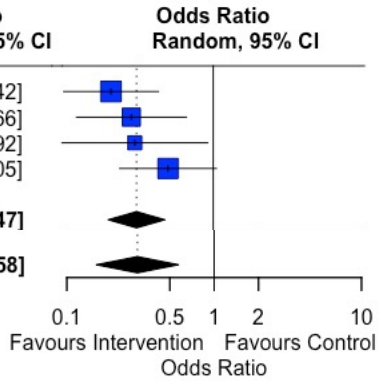
The **HKSJ** method is not always conservative compared to the **common-effect** meta-analysis!
This is why **the most conservative CI** is always **recommended** to be selected

HKSJ when the number of studies is <5

Number of studies : k=4

Study	Intervention Events	Total	Control Events	Total	Weight	Odds Ratio Random, 95% CI
Study A	9	345	40	342	31.3%	0.20 [0.10; 0.42]
Study C	7	286	24	290	24.1%	0.28 [0.12; 0.66]
Study D	4	200	13	200	14.4%	0.29 [0.09; 0.92]
Study E	12	116	22	116	30.1%	0.49 [0.23; 1.05]
Random effects model (WT)		**947**		**948**	**100.0%**	**0.30 [0.19; 0.47]**
Random effects model (HKSJ)		**947**		**948**	**100.0%**	**0.30 [0.16; 0.58]**

Heterogeneity: $Tau^2 = 0.0233$; $Chi^2 = 2.78$, df = 3 (P = 0.4270)

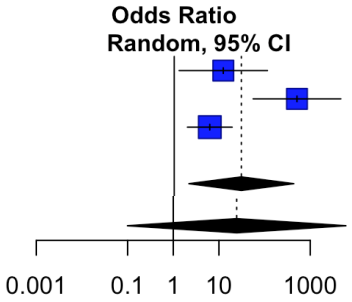


$\hat{\tau}^2 = 0.023$

Number of studies : k=3

Study	Intervention Events	Total	Control Events	Total	Weight	Odds Ratio Random, 95% CI
Rabe 1985	53	54	46	55	31.2%	10.37 [1.27; 84.97]
Christensen 1984	63	64	10	65	31.3%	346.50 [42.98; 2793.60]
Ho 1983	48	53	37	58	37.4%	5.45 [1.88; 15.81]
Random effects model (Wald Type)		**171**		**178**	**100.0%**	**24.49 [2.02; 297.47]**
Random effects model (HKSJ)						**24.49 [0.10; 6035.38]**

Heterogeneity: $Tau^2 = 4.0438$; $Chi^2 = 12.10$, df = 2 (P < 0.01)



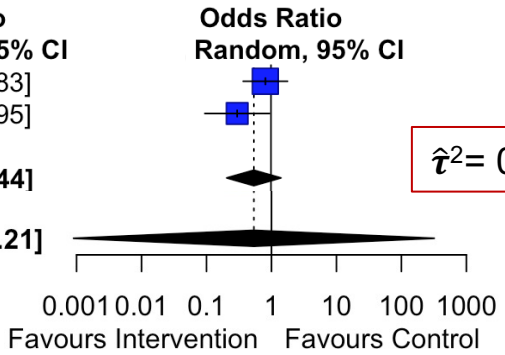
$\hat{\tau}^2 = 4.044$

Kapp et al CDSR 2010: https://pubmed.ncbi.nlm.nih.gov/20166091/

HKSJ when the number of studies is <5

Study	Intervention Events	Intervention Total	Control Events	Control Total	Weight	Odds Ratio Random, 95% CI
Caramez 1998	23	46	27	49	59.1%	0.81 [0.36; 1.83]
Silverman 2005	4	117	11	102	40.9%	0.29 [0.09; 0.95]
Random effects model (Wald Type)		**163**		**151**	**100.0%**	**0.54 [0.20; 1.44]**
Random effects model (HKSJ)						**0.54 [0.00; 320.21]**

Heterogeneity: $Tau^2 = 0.2585$; $Chi^2 = 1.98$, df = 1 (P = 0.16); $I^2 = 49\%$



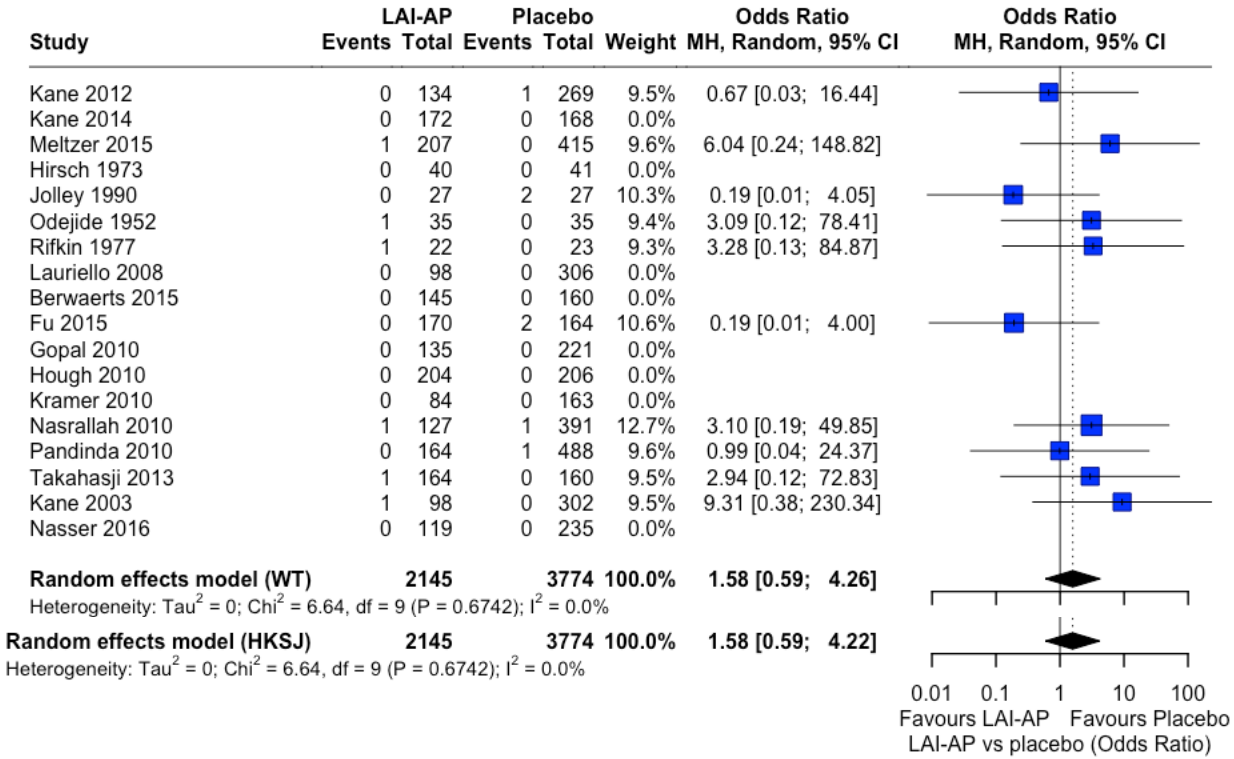
$\hat{\tau}^2 = 0.259$

In the case of 2 studies, the HKSJ can lead to overly conservative results!

Bain et al CDSR 2014: https://pubmed.ncbi.nlm.nih.gov/25331331/

In case of **rare events**, HKSJ performs worse than DL

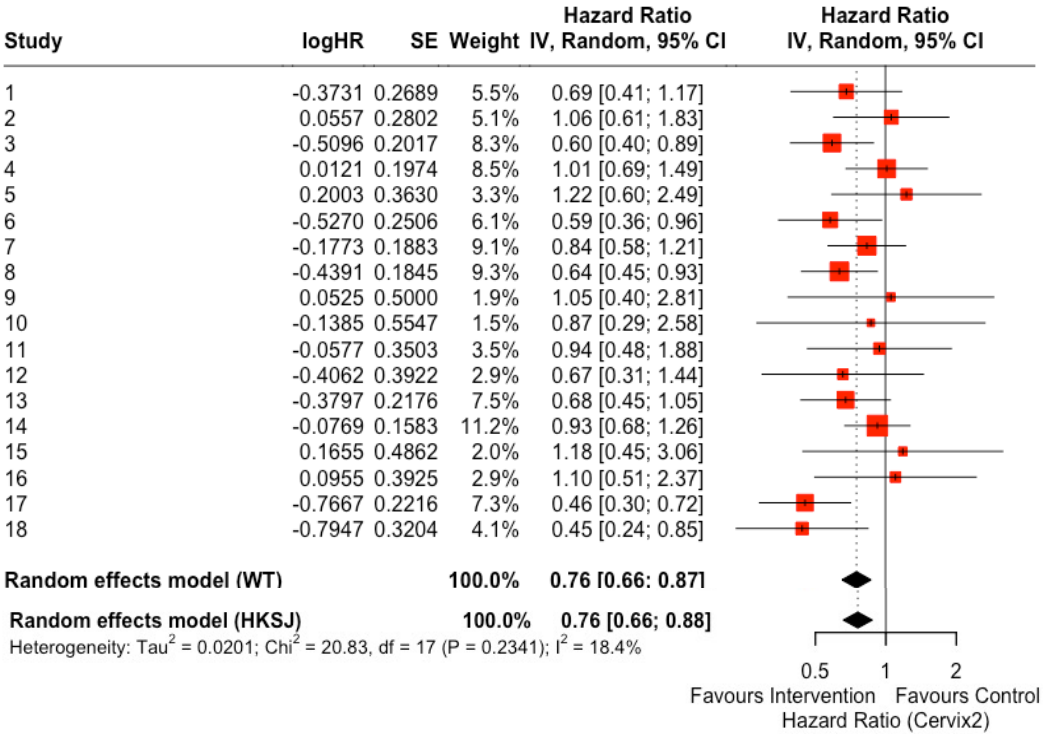
All-cause mortality in antipsychotics



$\hat{\tau}^2 = 0.00$

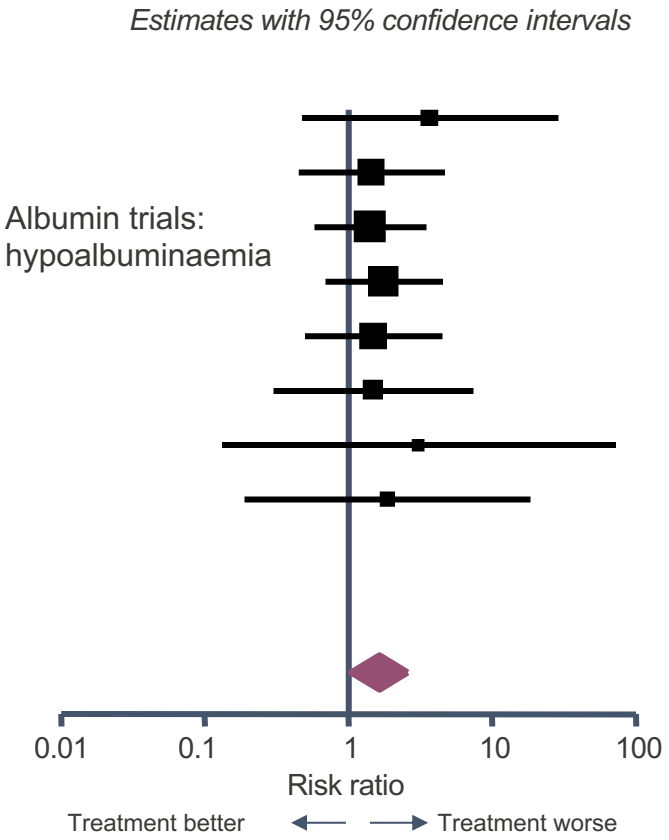
HKSJ is suboptimal than WT in meta-analyses with binary outcomes **of rare events !**

HKSJ gives comparable results to DL as the number of studies increases

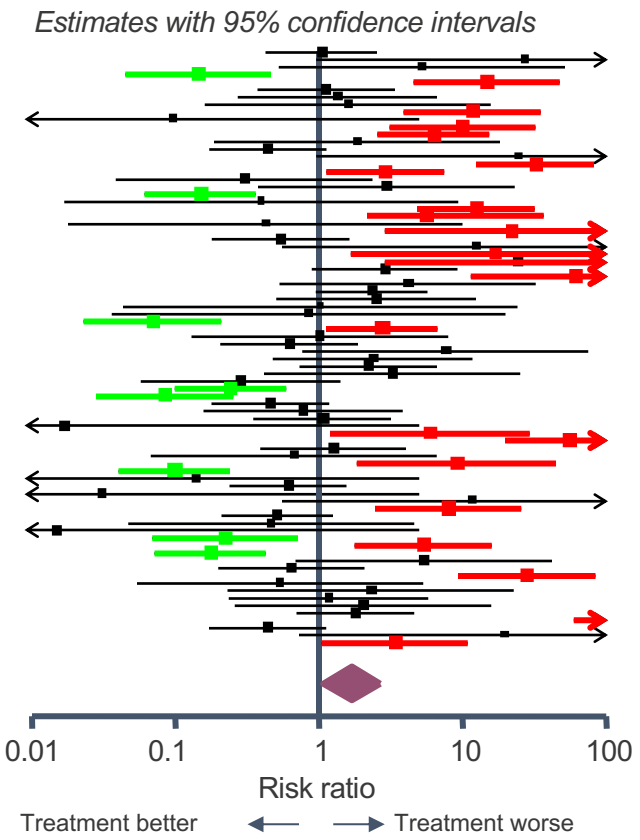


WT: [0.66, 0.87]

HKSJ: [0.66, 0.88]



Random effects meta-analysis:
1.64 (1.04 , 2.58) P = 0.03



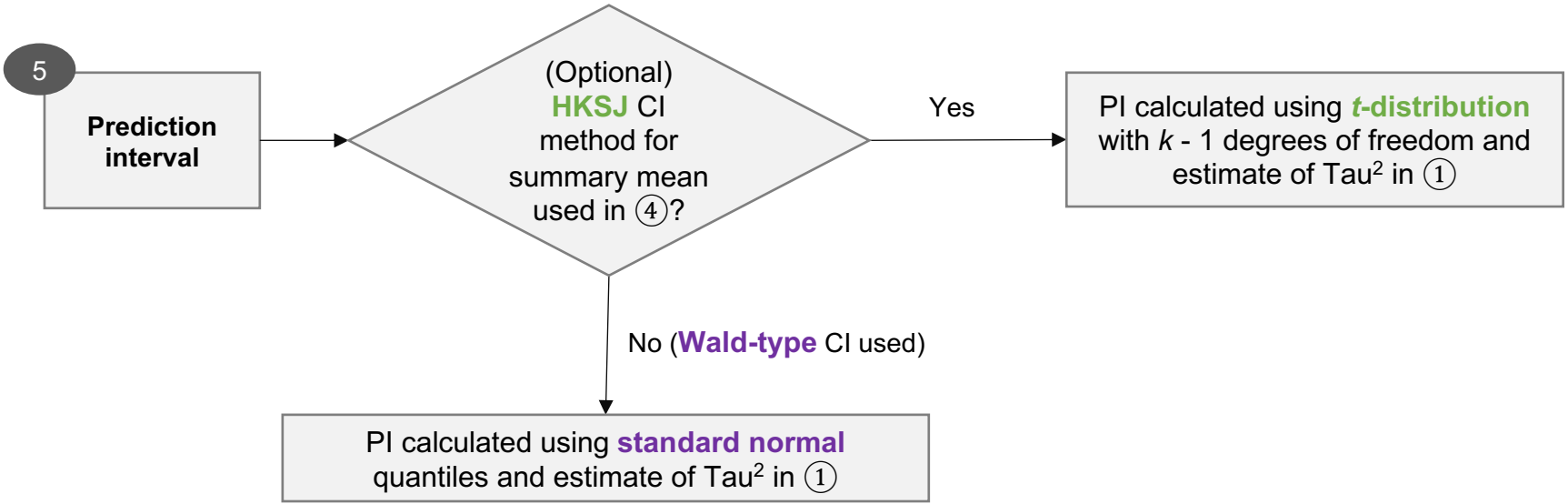
Random effects meta-analysis:
1.64 (1.04 , 2.58) P = 0.03

Cochrane

Source: Julian Higgins

Prediction intervals for random-effects meta-analysis

Quantity to be estimated	Statistical methods and decision rules implemented in RevMan to estimate the quantity



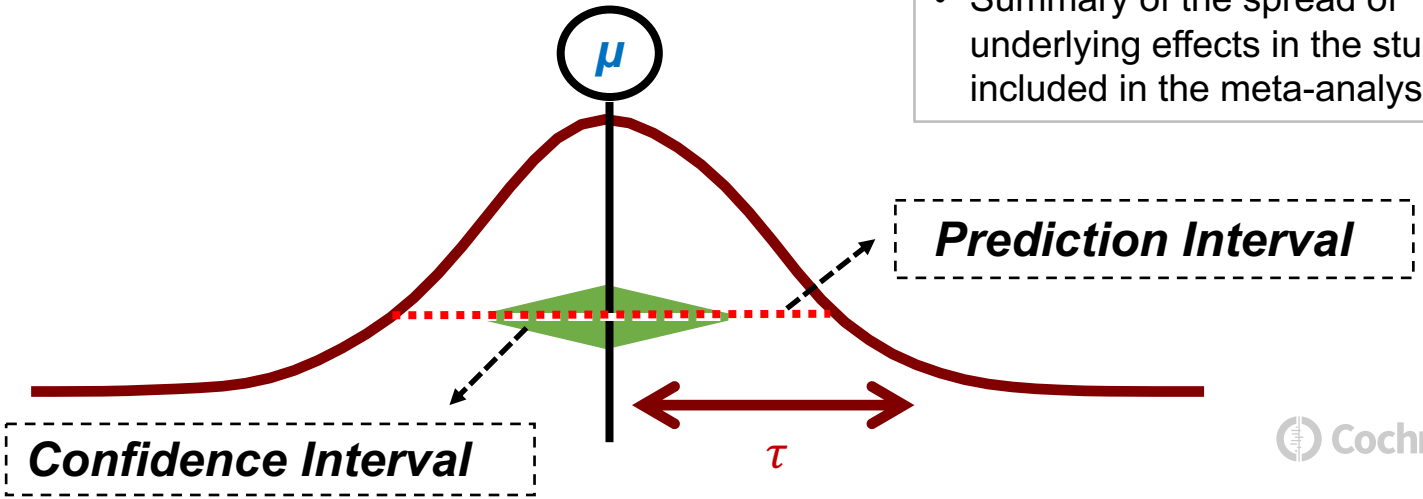
- CI = confidence interval, PI = prediction interval, REML = restricted maximum likelihood, *k* = number of studies,
- Tau^2 = estimated between-study variance, SE^2 = estimated 'typical' within study variance, HKSJ = Hartung-Knapp and Sidik-Jonkman

Prediction Intervals for random-effects meta-analysis

A 95% prediction interval where approximately 95% of the true treatment effects are predicted to fall is:

$$\hat{\mu} \pm 1.96\sqrt{\tau^2}$$

- The interval within which we expect that the effect of a future study will lie
- Summary of the spread of underlying effects in the studies included in the meta-analysis



Calculation of a prediction interval

An approximate 95% range of normally distributed underlying effects can by obtained by:

$$\hat{\mu}_{RE} \pm 1.96\sqrt{\tau^2}$$

But, in practice, both the summary estimate (μ) and τ are estimated, which needs to be accounted for when calculating the prediction interval:

$$\hat{\mu}_{RE} \pm m\sqrt{\hat{\tau}^2 + var(\hat{\mu}_{RE})}$$

Calculation of a prediction interval

An approximate 95% range of normally distributed underlying effects can by obtained by:

$$\hat{\mu}_{RE} \pm 1.96\sqrt{\tau^2}$$

But, in practice, both the summary estimate (μ) and τ are estimated, which needs to be accounted for when calculating the prediction interval:

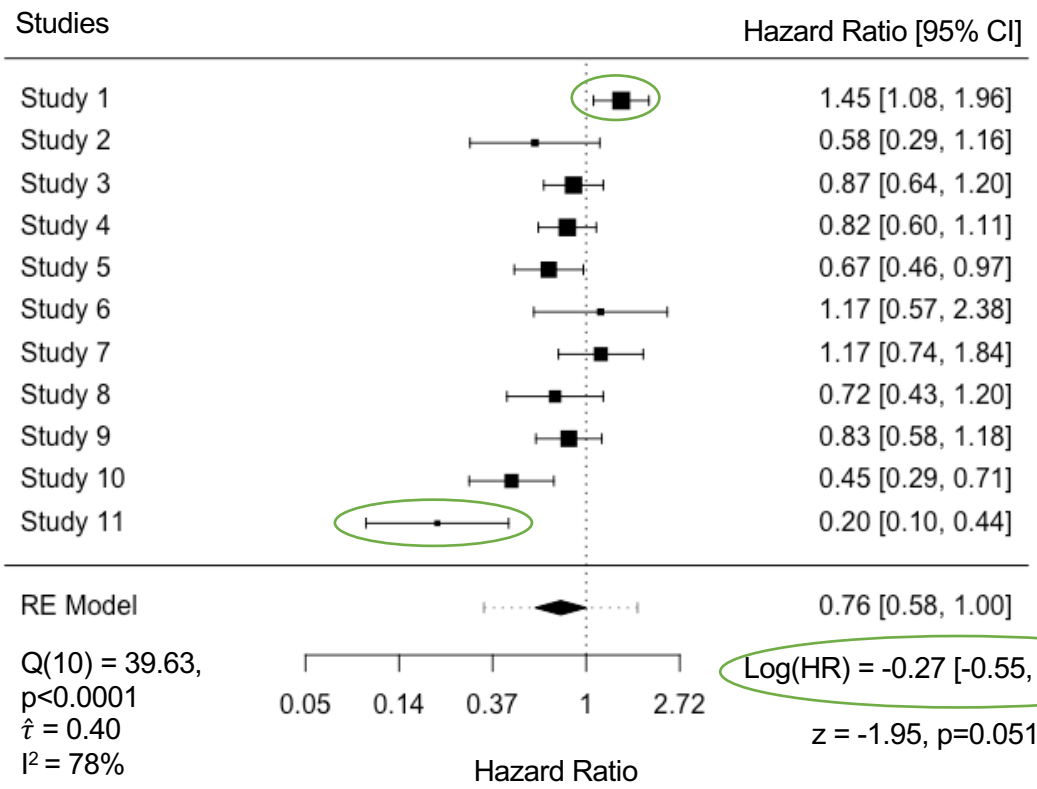
$$\hat{\mu}_{RE} \pm m\sqrt{\hat{\tau}^2 + var(\hat{\mu}_{RE})}$$

Choice of multiplier (m) is dependent on the confidence interval method used for the summary estimate

- Wald-type CI method → z quantile for PI
- HKSJ CI method → *t*-distribution with k-1 degrees of freedom

This choice of multiplier means that in the absence of observed heterogeneity, the CI and PI will be identical

Prediction Interval example



→ Log-transformed

Does the treatment reduce the risk of having an event? Does the effect size vary across studies?

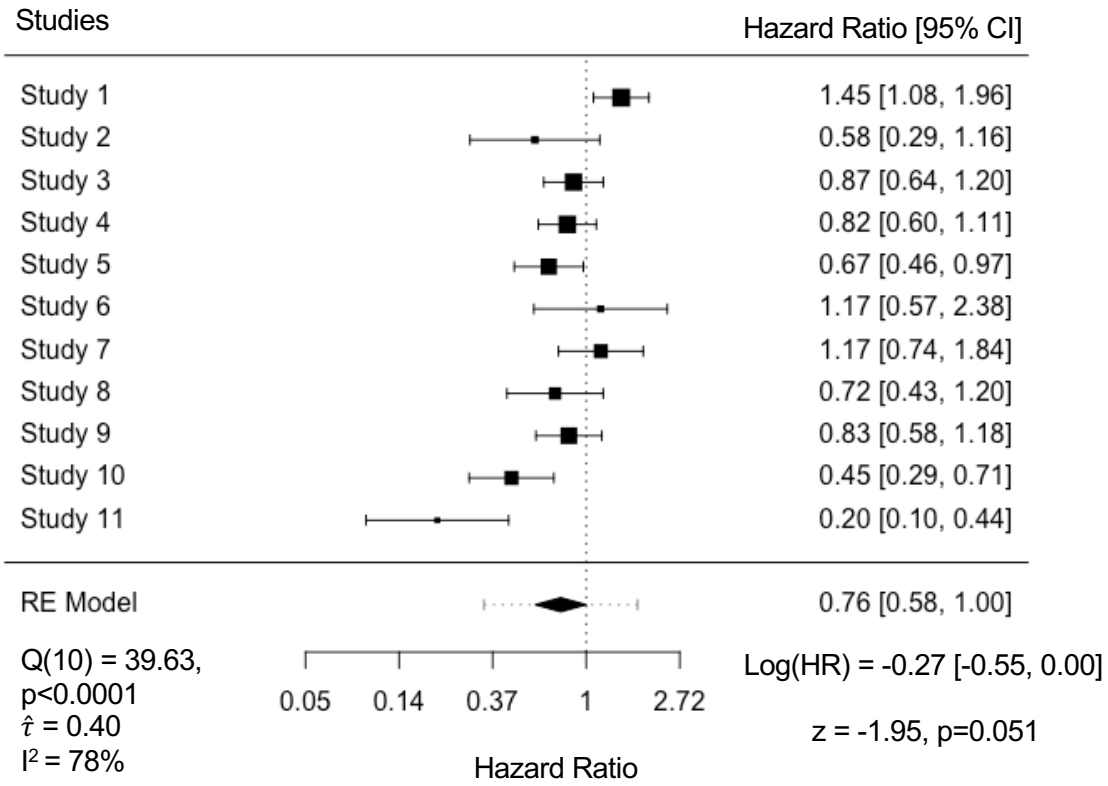
Calculation of Prediction Interval:

$$-0.27 \pm 1.96\sqrt{0.40^2 + 0.14^2}$$
$$[-1.10, 0.56]$$

Back-transformed Prediction Interval:

$$[0.33, 1.75]$$

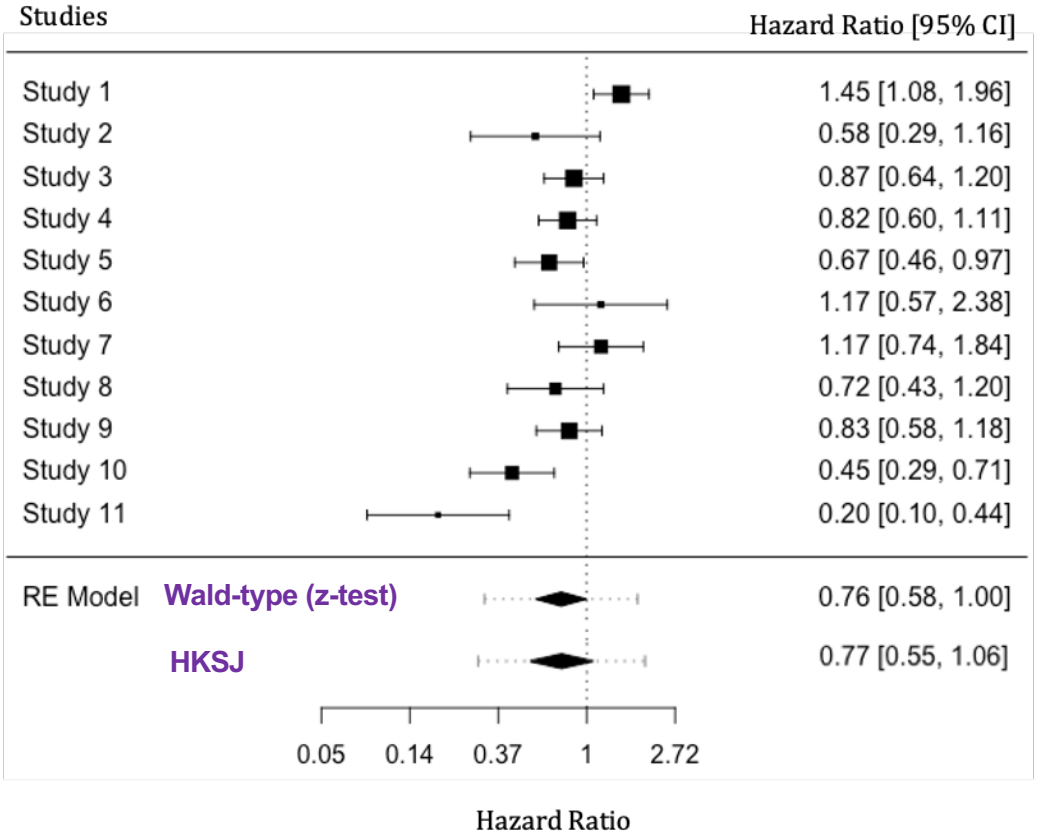
Prediction Interval example



Confidence Interval (precision):
The true summary logHR probably falls in the interval -0.55 to 0.00 (back-transformed to HR scale: 0.58 to 1.00)

Prediction Interval (dispersion):
The true logHR for any single study will probably fall in the range of -1.10 to 0.56
(back-transformed to HR scale: 0.33 to 1.75)

Prediction Interval example



[0.33, 1.75] **(normal distribution)**

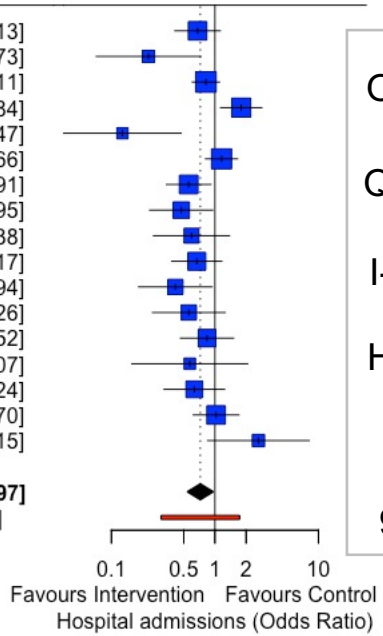
[0.32, 1.83] **(t_{k-1} distribution)**

Example using the recommended methods

Effect of quality improvement strategies for coordination of care on hospital admissions

Study	Intervention Events	Intervention Total	Control Events	Control Total	Weight	Odds Ratio MH, Random, 95% CI
Beck	35	160	47	161	7.4%	0.68 [0.41; 1.13]
Botha	13	32	18	24	3.2%	0.23 [0.07; 0.73]
Burns	210	353	228	355	9.1%	0.82 [0.60; 1.11]
Franklin	62	213	38	204	7.8%	1.79 [1.13; 2.84]
Lafave	13	24	37	41	2.7%	0.13 [0.03; 0.47]
Puschner	108	241	103	250	8.7%	1.16 [0.81; 1.66]
Rich (a)	41	142	59	140	7.5%	0.56 [0.34; 0.91]
Salkever	27	91	25	53	5.8%	0.47 [0.23; 0.95]
Rich (b)	21	63	16	35	4.8%	0.59 [0.25; 1.38]
Kasper	47	102	55	98	6.9%	0.67 [0.38; 1.17]
Courtney	13	49	27	58	5.0%	0.41 [0.18; 0.94]
Castro	20	50	25	46	5.0%	0.56 [0.25; 1.26]
Burns	17	110	56	313	6.6%	0.84 [0.46; 1.52]
Koehler	6	20	9	21	2.8%	0.57 [0.16; 2.07]
Ruchlewska	24	70	33	73	6.0%	0.63 [0.32; 1.24]
Laramee	49	131	46	125	7.4%	1.03 [0.62; 1.70]
Stewart	24	29	31	48	3.3%	2.63 [0.85; 8.15]
Random effects model (REML)		**1880**		**2045**	**100.0%**	**0.72 [0.54; 0.97]**
Prediction interval						**[0.30; 1.73]**

Heterogeneity: $Tau^2 = 0.1539$; $Chi^2 = 42.66$, df = 16 (P = 0.0003); $I^2 = 62.5\%$



OR: 0.72 with a 95% CI [0.54, 0.97]

Q-statistic: 42.66, df: 16, P<0.001

I-squared: 62.5%

Heterogeneity variance: $\hat{\tau}^2 = 0.154$

95% CI for $\hat{\tau}^2$: [0.052, 0.838]

95% prediction Interval: [0.30, 1.73]

Tricco et al CMAJ 2014

Odds ratios less than 1.0 indicate decreased odds of admission to hospital

What to write in a protocol?

PRISMA 2020 – item 13d

Describe any methods used to synthesise results and provide a rationale for the choice(s). If meta-analysis was performed, **describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used**

Essential elements (some):

If meta-analysis was done, specify:

- the meta-analysis model (fixed-effect, fixed effects, or random-effects) and provide rationale for the selected model
- the method used (such as Mantel-Haenszel, inverse-variance)
- any methods used to identify or quantify statistical heterogeneity (such as visual inspection of results, a formal statistical test for heterogeneity, heterogeneity variance (τ^2), inconsistency (such as I^2), and prediction intervals

What to write in a protocol?

PRISMA 2020 – item 13d (continued)

Describe any methods used to synthesise results and provide a rationale for the choice(s). If meta-analysis was performed, **describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used**

Essential elements (some):

If a random-effects meta-analysis model was used, specify:

- the between-study (heterogeneity) variance estimator used (such as DerSimonian and Laird, restricted maximum likelihood (REML))
- the method used to calculate the confidence interval for the summary effect (such as Wald-type confidence interval, Hartung-Knapp-Sidik-Jonkman)

What's next?

Demonstration of new random-effects methods in RevMan

There are many methods available to fit random-effects meta-analysis. However, until 2024, the only option available in RevMan has been the DerSimonian and Laird random-effects method. This method is known to have poor statistical performance in meta-analyses with characteristics commonly found in Cochrane reviews (e.g., meta-analyses with few studies). To address this issue, Cochrane is implementing new random-effects methods in RevMan. These include a new method for estimating the between-study (heterogeneity) variance, calculating the confidence interval for the summary effect, and adding prediction intervals to aid in interpreting random-effects meta-analysis findings.

In two web clinics, the presenters will provide participants with knowledge about these new methods and their implementation in RevMan. Specifically, in the first web clinic, the presenters will outline the new methods, while in this second clinic, they will demonstrate applying the new random-effects methods using RevMan.

Presenter Bios

Professor Jo McKenzie is head of the Methods in Evidence Synthesis Unit within the School of Public Health and Preventive Medicine at Monash University, Melbourne, Australia. She is Co-Convenor of the Cochrane Statistical Methods Group and an author of several chapters of the Cochrane Handbook for Systematic Reviews of Interventions.

Dr. Areti Angeliki Veroniki is a Scientist at the Knowledge Translation Program of the Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, and an Assistant Professor at the University of Toronto in the Institute of Health Policy, Management, and Evaluation. She is a Co-Convenor of the Cochrane Statistical Methods Group and Co-Chair of the Cochrane Methods Executive.

Sign up

Thursday, 12 December 2024, 20:00 UTC [check the time in your timezone] **SIGN UP HERE**

THANK YOU

Cochrane Methods
Statistics

a.veroniki@utoronto.ca

joanne.mckenzie@monash.edu