Cochrane

Changes to RevMan in terms of random-effects methods

Areti-Angeliki Veroniki, PhD, MSc

St. Michael's Hospital, Unity Health Toronto

IHPME, University of Toronto

Trusted evidence.
Informed decisions.
Better health.

Cochrane Methods
Statistics

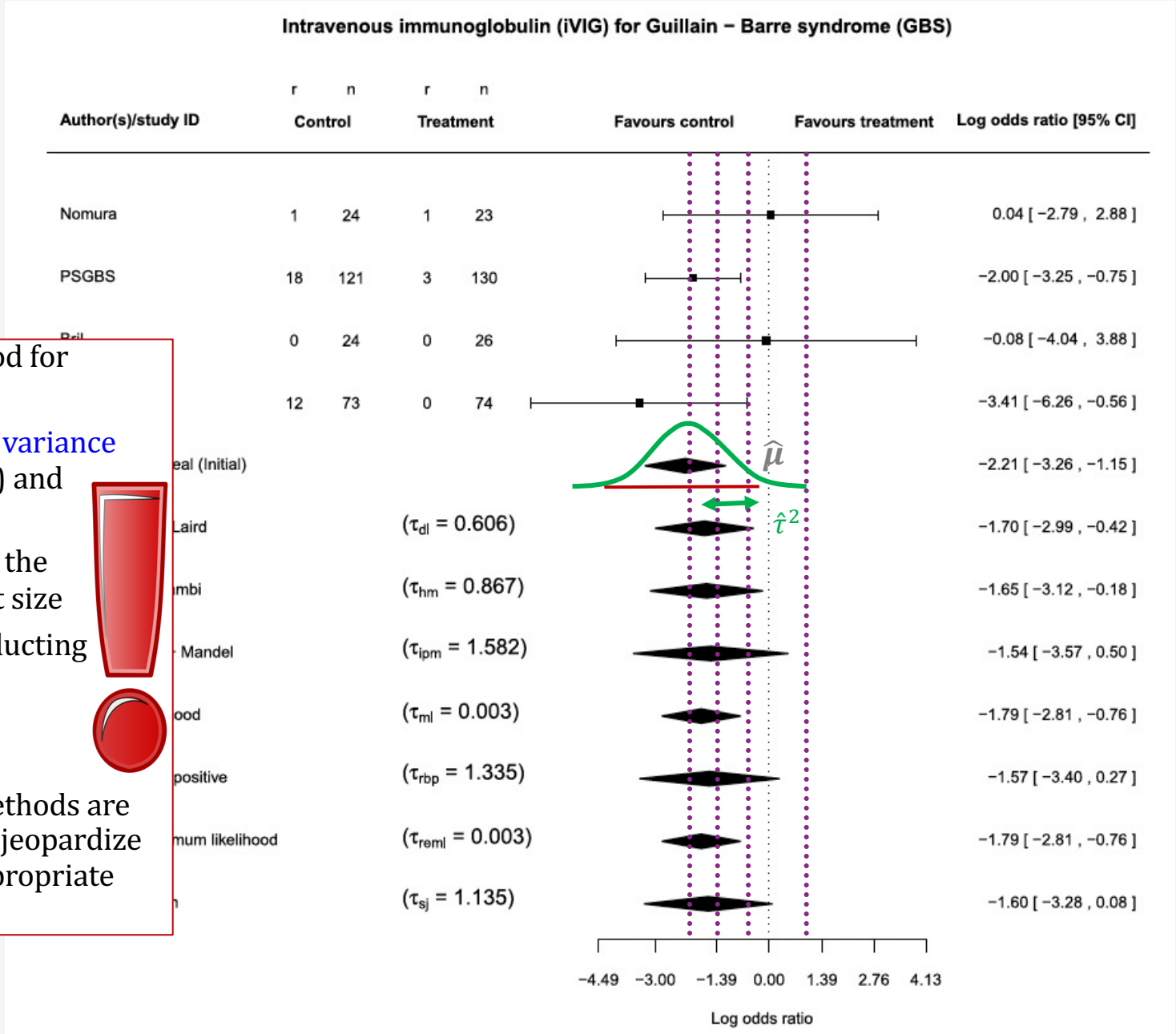
I have no actual or potential conflicts of interest in relation to this presentation

Acknowledgements

- Dean Langan
- Simon Turner
- Mark Simmonds*
- Anna Chaimani*
- Kerry Dwan
- Jo McKenzie*

* Denotes Convenors of the Cochrane Statistical Methods Group

Cochrane



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)

Cochrane

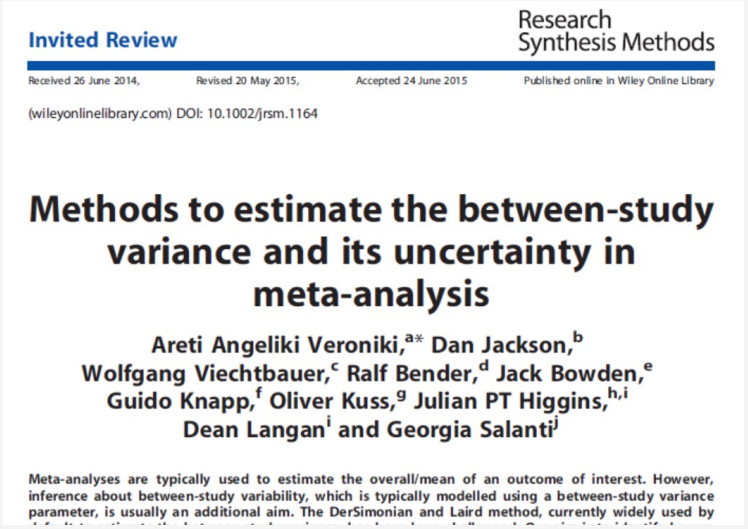
Hughes R. Cochrane Database Syst Rev 2014
Pateras et al Contemporary Clinical Trials Communications 2018

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

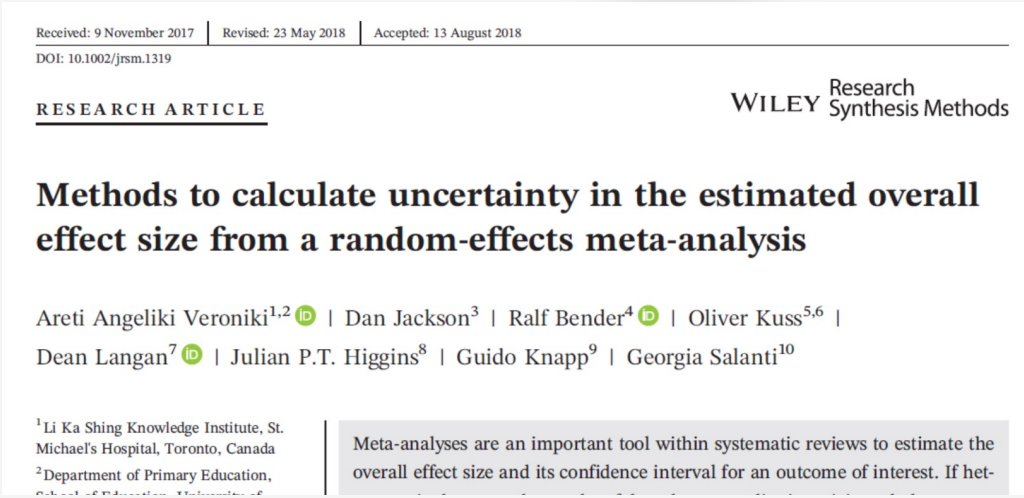
Received 26 June 2014, Revised 20 May 2015, Accepted 24 June 2015 Published online in Wiley Online Library

(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-

Process used by the SMG to develop recommendations

Step 1
Update of a systematic reviews of statistical simulation studies examining the performance of heterogeneity estimators and CI methods for summary effect size [Langan 2017, Veroniki 2016 and 2019]

Step 2
- Multiple meetings to review and discuss the evidence
- Form recommendations

Step 3
Examine the impact of adopting the recommendations when applied to meta-analyses in the Cochrane Library

Step 4
- Submit recommendations and evidence to the Cochrane Scientific Committee
- 'Recommendations approved by Cochrane Editor-in-Chief July 2022

Step 5
- 'Collaborate with RevMan Web developers and Cochrane Training to implement the methods and develop training

Team members:
- Areti Angeliki Veroniki
- Dean Langan
- Simon Turner
- Mark Simmonds
- Anna Chaimani
- Kerry Dwan
- Joanne McKenzie

Experience
- Co-convenors of the Cochrane Statistical Methods Group
- Led systematic reviews of statistical simulation studies, and undertaken simulation studies, examining random-effects methods
- Cochrane Methods Support Lead and Statistical Editor

Recommendations based on published studies

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

- Number and size of studies included in the meta-analysis
- Magnitude of true heterogeneity
- Distribution of true treatment effects
- Type of data (e.g., dichotomous, continuous)
- Choice of effect measure
- Frequency of events (for dichotomous outcomes)
- How well study-specific weights, variances and treatment effects are estimated – we often assume these are known.

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]

Cochrane

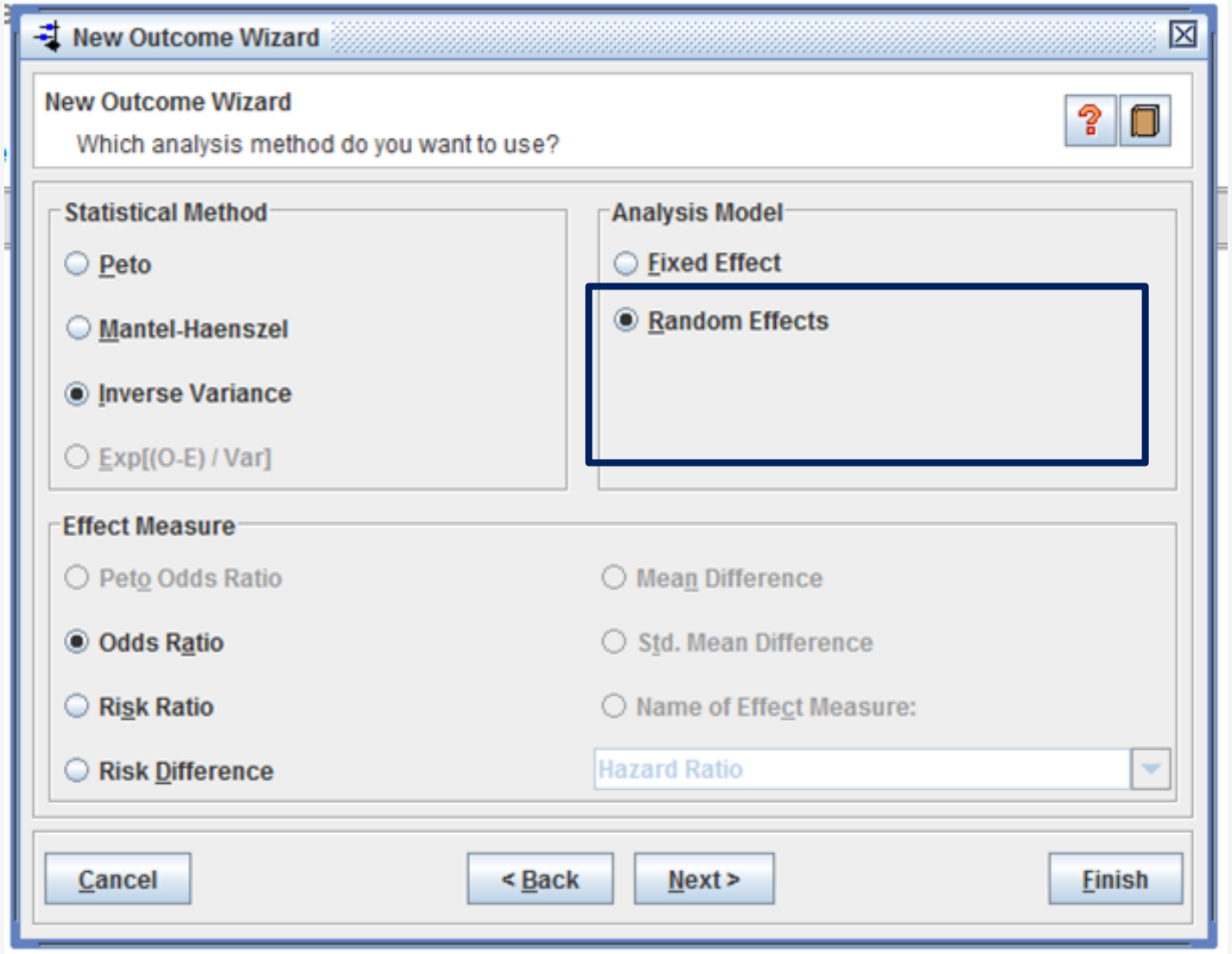
Summary of the properties of REML, PM and DL

			Number of studies	
			Few (2 - 6)	**Many (6+)**
Outcome type	**Continuous (mean difference and standardised mean difference)**		All estimators (REML, PM, DL) negatively biased in meta-analyses with moderate to high heterogeneity, particularly in meta-analyses with small studies. REML and PM biased to a lesser extent, REML recommended.	REML recommended, though performance is broadly comparable with PM and DL. PM not recommended in meta-analyses with large differences in study size. Otherwise, all methods perform reasonably well, particularly in meta-analyses with large studies.
	Dichotomous (odds ratio, relative risk, risk difference)	**$0.1 < p < 0.9$**	Performance of methods broadly comparable with continuous outcome meta-analyses (see above).	
		$p < 0.1$	All have considerable negative bias except for meta-analyses with large studies (where REML is recommended with better performance than DL).	All estimators (REML, PM, DL) negatively biased in meta-analyses with moderate to high heterogeneity, particularly in meta-analyses with small studies. REML and PM often biased to a lesser extent, REML recommended.
		$p < 0.01$	All estimators have substantial negative bias (REML, PM and DL) in meta-analyses with small study sizes. REML recommended if any, but all methods poor.	All have substantial negative bias except for meta-analyses with many large studies (where REML is recommended with better performance than DL)

** p in this table refers to the average event probability across the studies in a meta-analysis*

Cochrane

Heterogeneity variance estimators to be added to RevMan



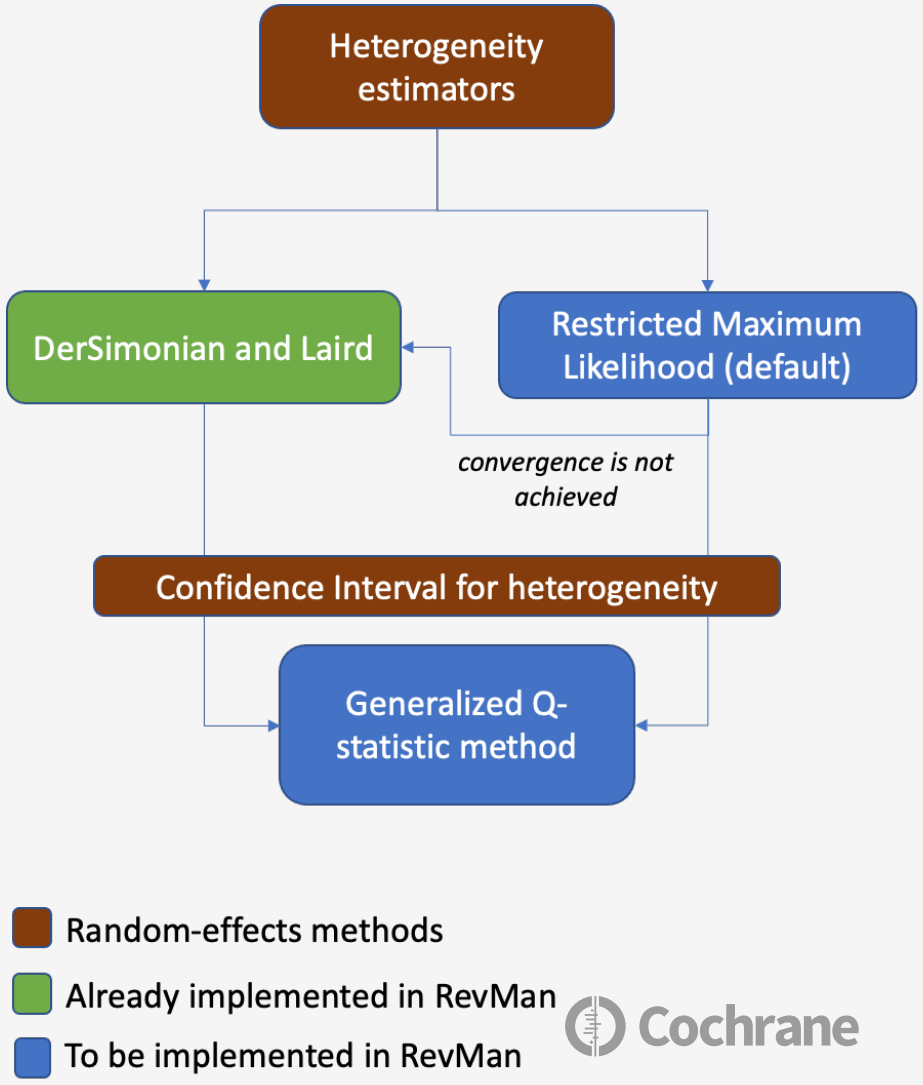
RevMan

SMG Recommendations

1. We recommend that **REML** as the default option

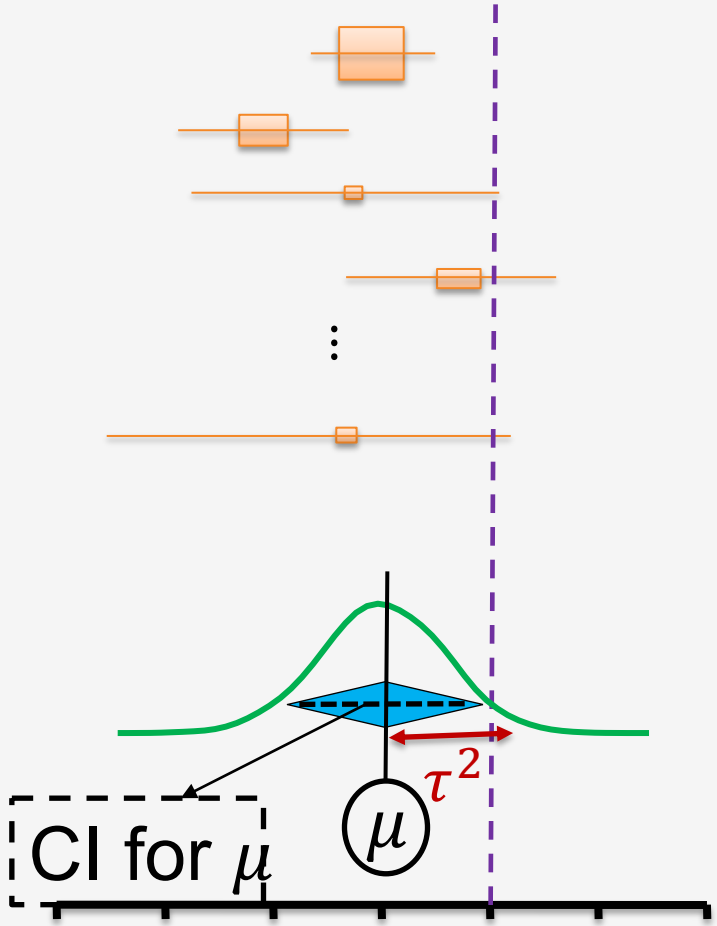
   – When REML cannot provide a unique estimate for heterogeneity variance (i.e., when a scoring algorithm, e.g., Fisher's scoring and Newton-Raphson, cannot solve the ML equations numerically), the DL method should be used as a closed form, non-iterative method

2. Methods to calculate CIs for the heterogeneity variance should be available in RevMan - We recommend the **generalized Q-statistic method**



Confidence Intervals (CIs) for summary effect

- There are >15 approaches to calculate a CI for the overall effect size under the random-effects model
  - Wald-type method (some refer to this method as DerSimonian and Laird) is the most popular CI approach
- Hartung-Knapp-Sidik-Jonkman method
  - Choice of τ^2 estimator becomes much less important
  - But, with 2 to 3 studies, CI can be very wide
  - In the absence of heterogeneity: HKSJ coverage < WTz coverage
  - Making inferences in the random-effects model with a *small* number of studies is **hard**!



Cochrane

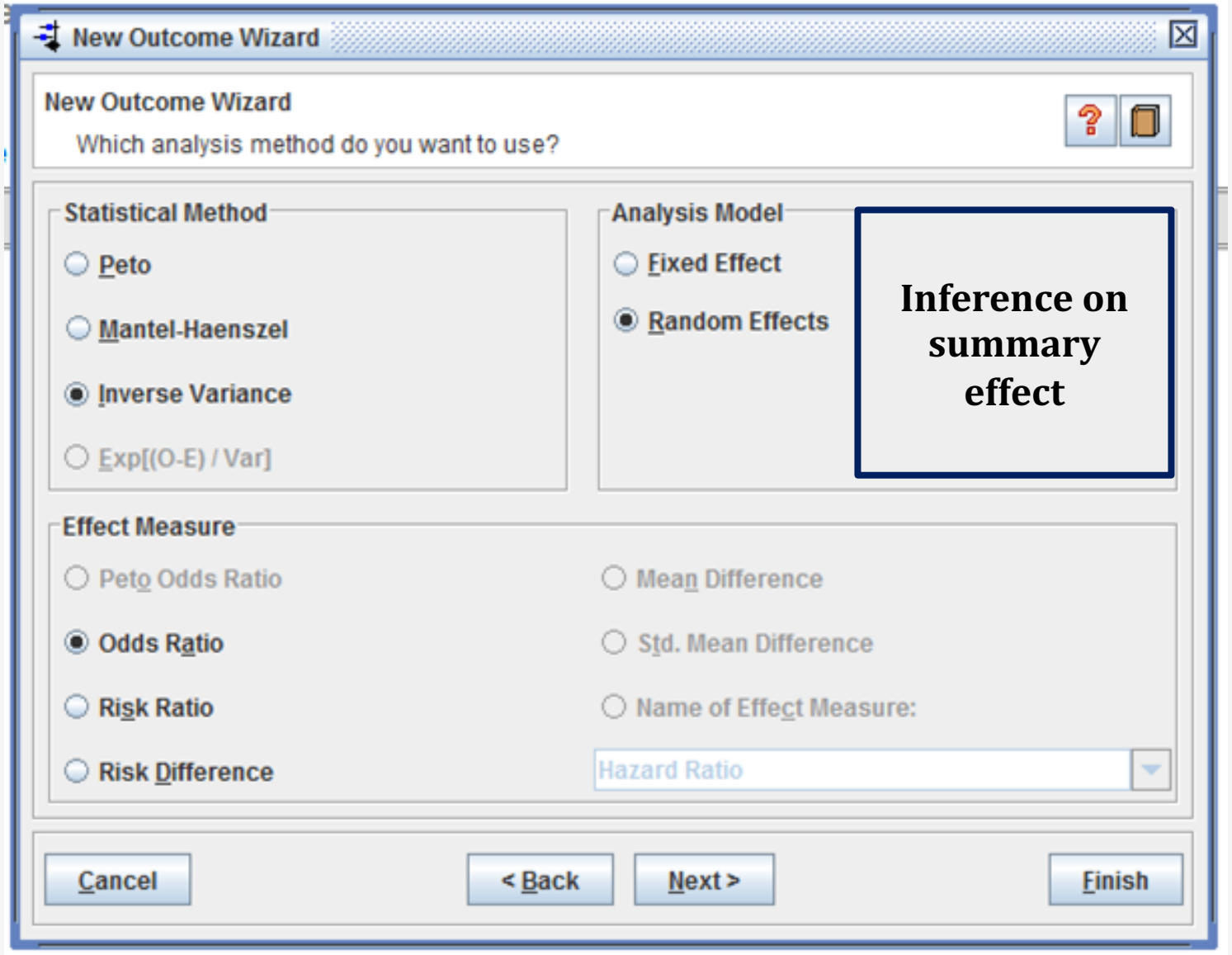
Summary of the properties of Wald type (z-test), HKSJ, and modified HKSJ

		Number of studies		
		Small (2)	**Small (3 – 4)**	**Moderate to Large (≥5)**
Presence of observed heterogeneity	**Yes**	Both Wald type and HKSJ have poor coverage.	The modified HKSJ outperforms HKSJ and Wald type methods in terms of coverage.	HKSJ outperforms the Wald type method in terms of coverage. The modified HKSJ leads to overly conservative results compared with HKSJ.
	No	The Wald type method outperforms HKSJ with higher power (e.g., HKSJ: 15% vs Wald type: 60%).	The Wald Type outperforms HKSJ in terms of coverage.	The Wald Type outperforms HKSJ in terms of coverage.

** modified HKSJ with q*=max{q, 1}*

Cochrane

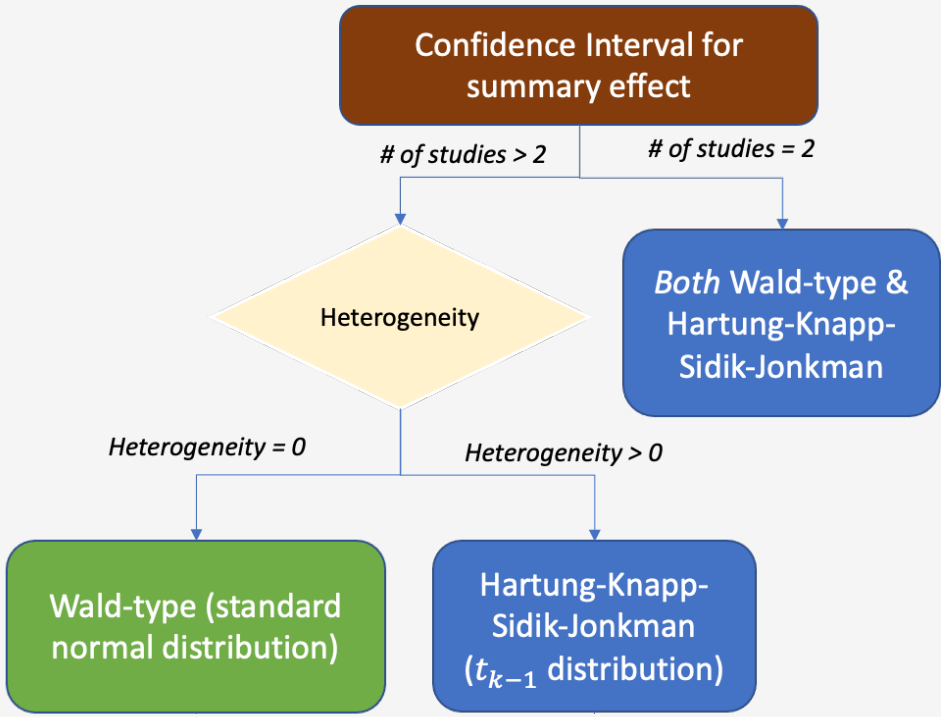
Confidence Interval methods to be added to RevMan



RevMan

SMG Recommendations

1. We recommend that **HKSJ** as the default option when:
   - the number of studies in the meta-analysis is >2, and
   - the estimated heterogeneity variance is >0
2. But, the **Wald-type (z-test)** in the absence of heterogeneity
3. For meta-analyses with 2 studies, we recommend the use of both **HKSJ** and **Wald-type (z-test)** methods



Cochrane

Interpreting random effects meta-analysis

Random-effects meta-analysis suitable for unexplained heterogeneity

- Random effects may not explain all the heterogeneity of the data if covariates are responsible

Conventionally, inference is focused on the mean of the distribution ($\hat{\mu}$)

- i.e. we report mean and 95% CI (measure of precision) from a meta-analysis
- What about the dispersion of the effect size?

Can also calculate a prediction interval

$$\hat{\mu}_{RE} \pm z_{0.975}\sqrt{\hat{\tau}^2 + var(\hat{\mu}_{RE})}$$

$$\hat{\mu}_{RE} \pm t_{k-1,0.975}\sqrt{\hat{\tau}^2 + var(\hat{\mu}_{RE})}$$

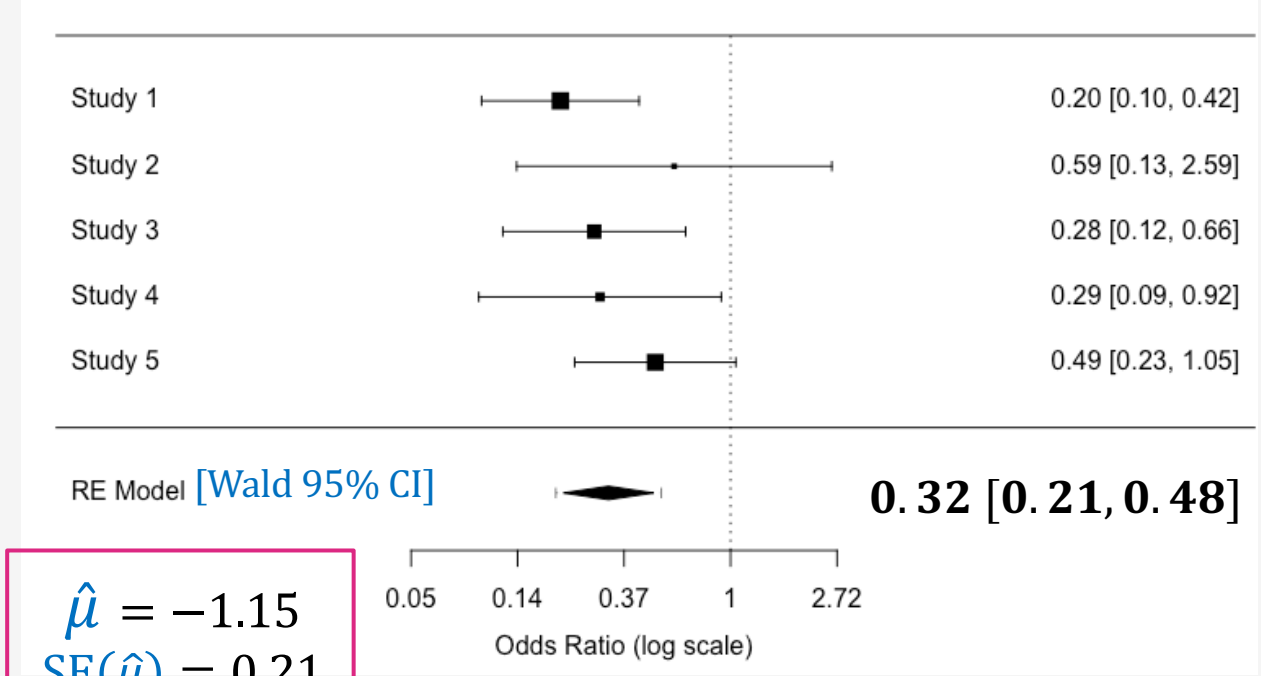
The interval within which we expect that the effect of a future study will lie

- 95% of the true effect estimates are predicted to fall within the aforementioned interval

Borenstein et al RSM (2016)

Cochrane

Prediction Intervals (PIs)



$\hat{\mu} = -1.15$
$\mathrm{SE}(\hat{\mu}) = 0.21$

95% PI: $\hat{\mu} \pm z_{0.975}\sqrt{\hat{\tau}^2 + var(\hat{\mu})}$

$$\exp\left(-1.15 \pm 1.96\sqrt{0.0175 + 0.21^2}\right)$$

[0.19, 0.52]

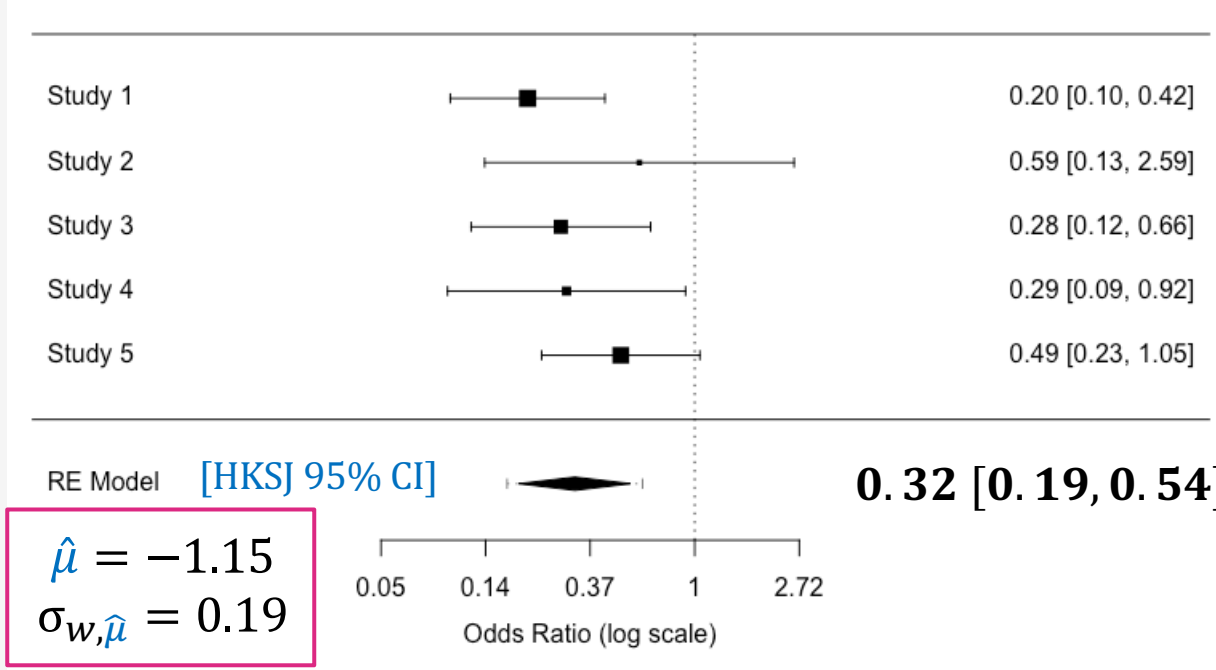
95% PI: $\hat{\mu} \pm t_{k-1,0.975}\sqrt{\hat{\tau}^2 + \sigma^2_{w,\hat{\mu}}}$

$$\exp\left(-1.15 \pm 2.78\sqrt{0.0175 + 0.19^2}\right)$$

[0.17, 0.61]

$\sigma^2_{w,\hat{\mu}} = q \cdot var(\hat{\mu})$

$q = \frac{Q_{gen}}{k-1}$, and $Q_{gen} = \sum w_{i,RE}(y_i - \hat{\mu})^2$



$\hat{\mu} = -1.15$
$\sigma_{w,\hat{\mu}} = 0.19$

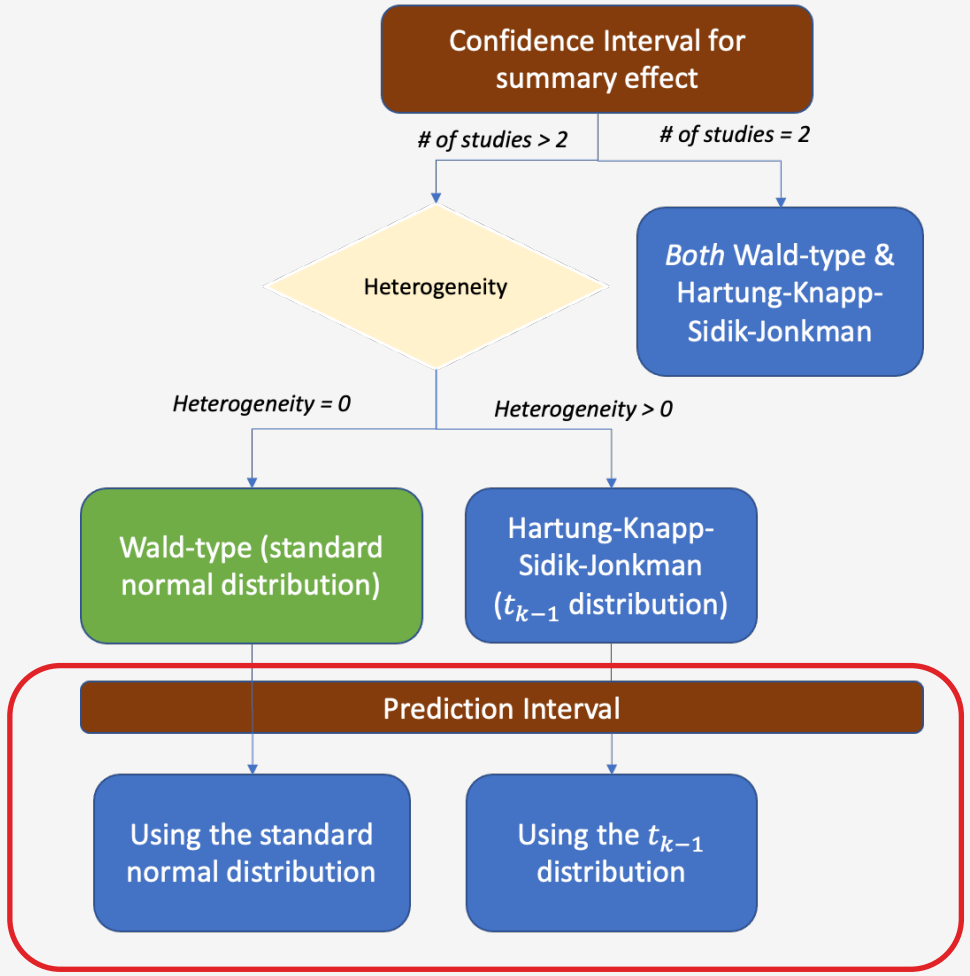
SMG Recommendations

1. Methods to calculate prediction intervals **should be** implemented

2. We recommend:

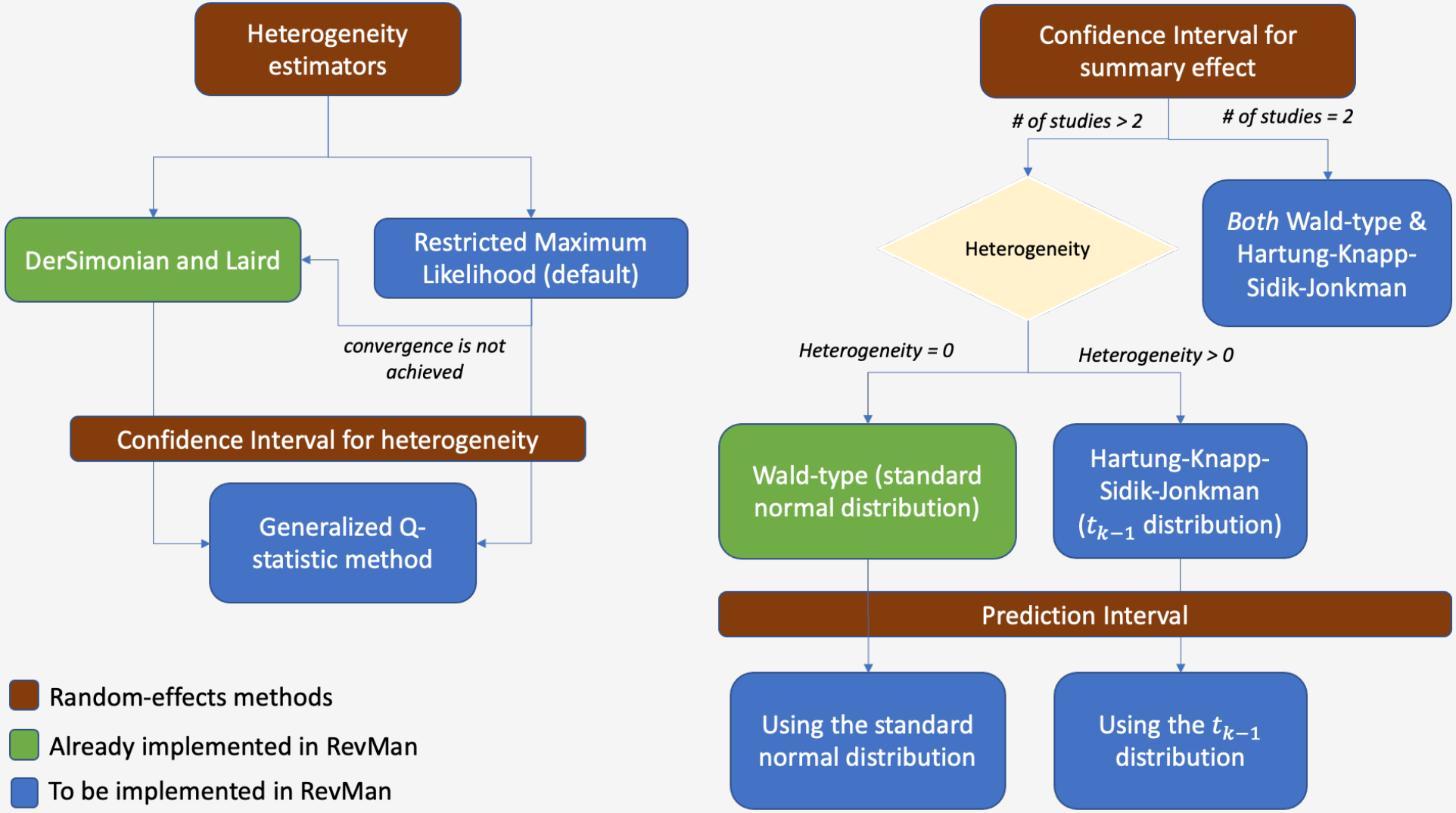
   - The **t-distribution prediction interval** when the **HKSJ** is used

   - The **normal distribution prediction interval** when the **Wald-type normal distribution** is used



Cochrane

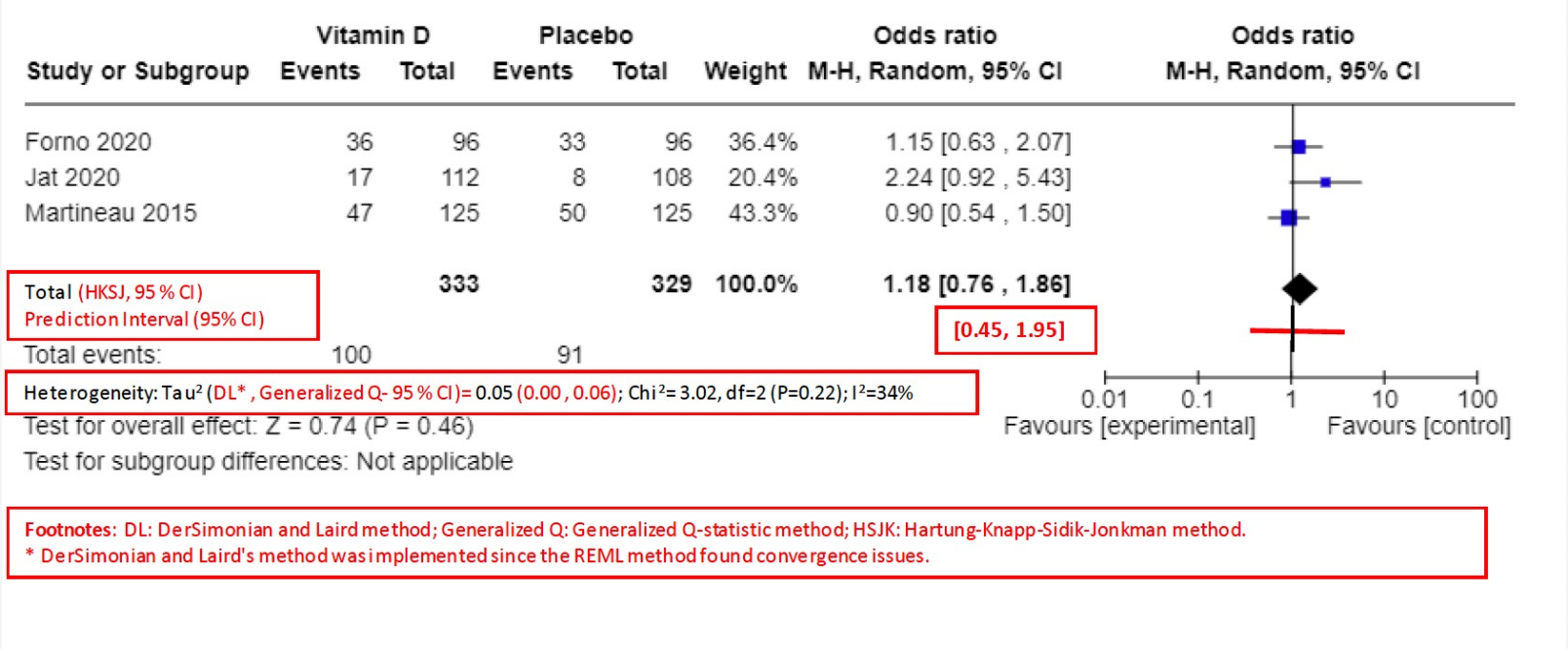
SMG Overall Recommendations



Time for CHANGE!

Next Steps

- There are still some decisions to be made
  - e.g. reporting two confidence intervals using different methods for meta-analyses with 2 studies
- We hope to make these features available during Q4-2023
  - During the transition period (~6 months) these changes will become optional for authors and mandatory after ~ 6 months
- After the transition period these changes will become mandatory for authors (except for the calculation of prediction intervals)



Cochrane

@AVeroniki

a.veroniki@utoronto.ca

Cochrane