

# **Pairwise meta-analysis**

**Georgia Salanti**

Institute of Social and Preventive Medicine  
University of Bern

# Outline

- Systematic reviews and meta-analysis
- Approaches to meta-analysis
- Fixed-effect(s) meta-analysis
- Random-effects meta-analysis
- Heterogeneity
- Fixed vs random-effects meta-analysis
- [pollev.com/gmhbe](http://pollev.com/gmhbe)

# NEWS

Home | UK | World | Business | Politics | Tech | Science | Health | Education | Entertainment

## Health

### NHS screening plan for type 2 diabetes 'inaccurate'

© 5 January 2017 | Health |

Share

They also found that lifestyle interventions lasting three to six years showed a 37% reduction in relative risk of type 2 diabetes - equivalent to 90 fewer people in every 1,000 developing the disease.



SCIENCE PHOTO LIBRARY

The NHS programme for screening those at high risk of type 2 diabetes is unlikely to have much impact, an Oxford University study in the BMJ suggests.

It concluded that inaccurate blood tests would give too many people an incorrect diagnosis, while lifestyle changes had a low success rate.

# Efficacy and effectiveness of screen and treat policies in prevention of type 2 diabetes: systematic review and meta-analysis of screening tests and interventions

Eleanor Barry,<sup>1</sup> Sam...  
Trisha Greenhalgh

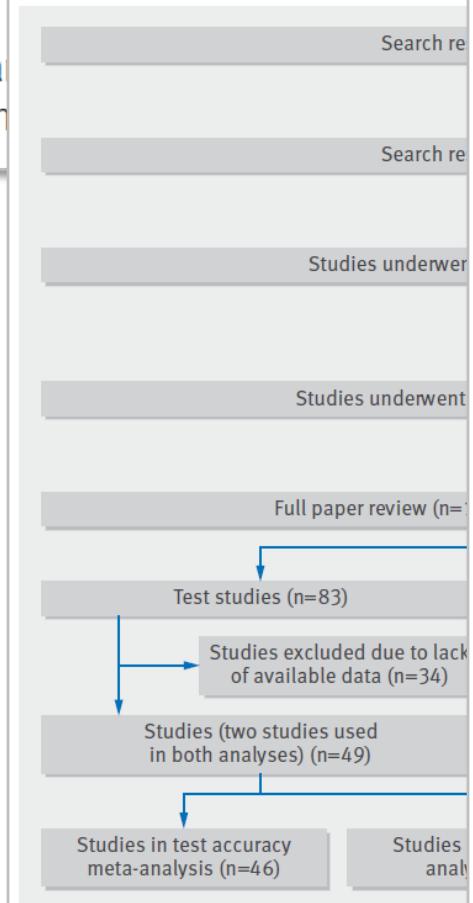


Fig 1 | Flow diagram of studies identified and included in the systematic review of the effectiveness of screen and treat policies in prevention of type 2 diabetes.

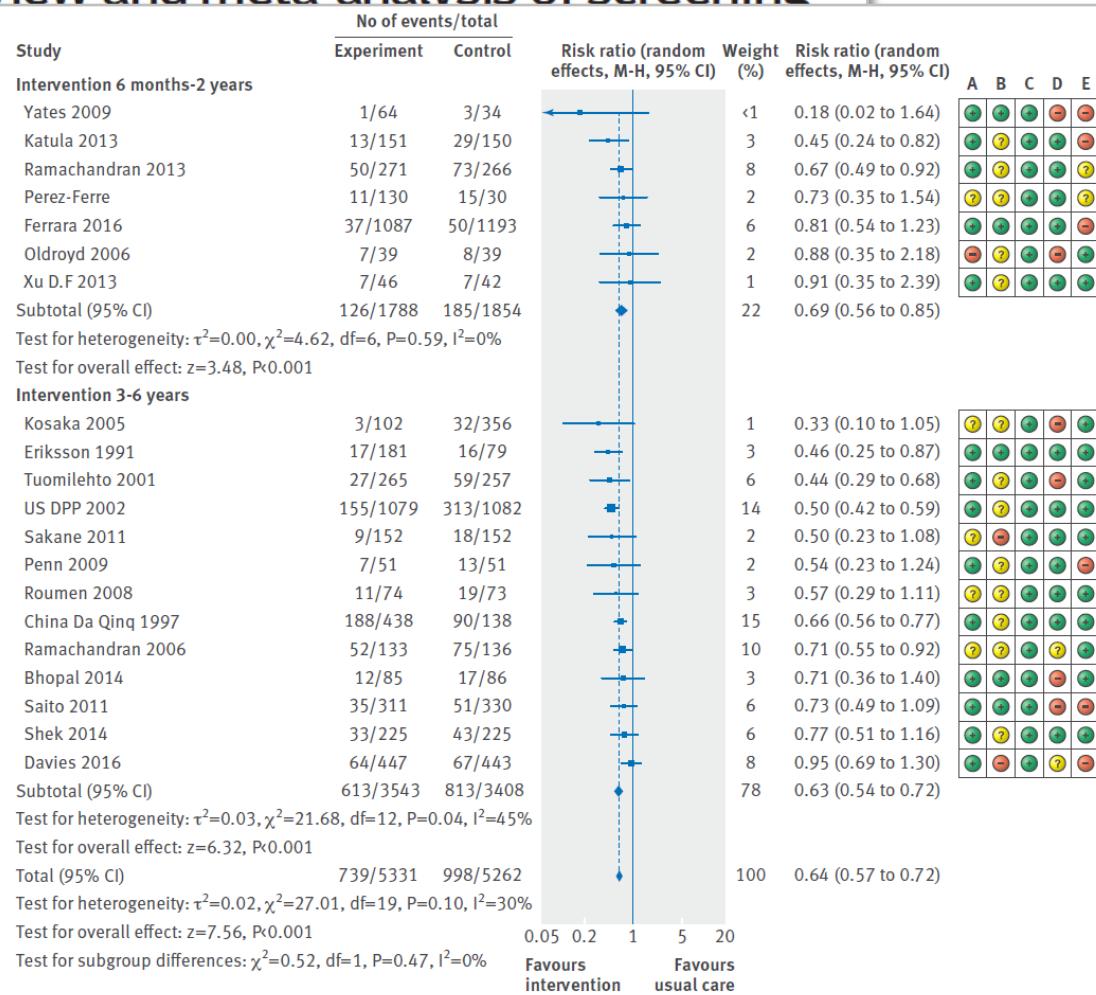
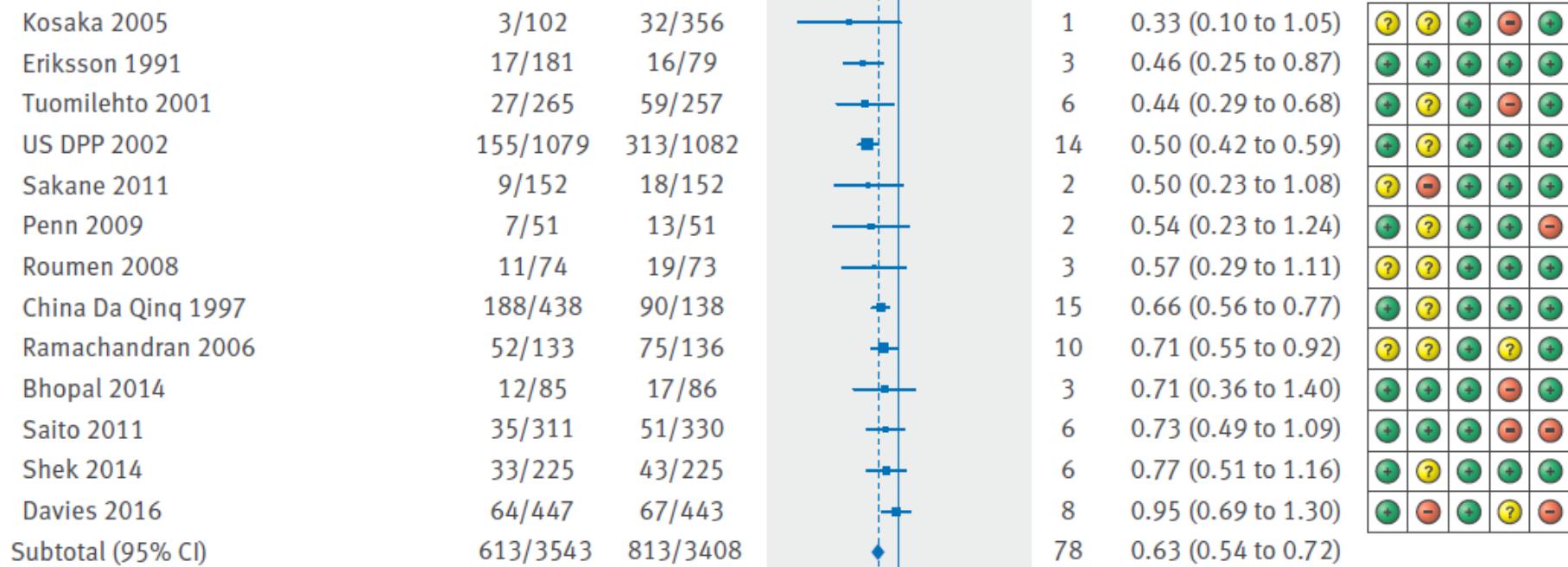


Fig 6 | Relative reduction in risk of diabetes at end of lifestyle trials. A=random sequence generation (selection bias); B=allocation concealment (selection bias); C=blinding of outcome assessment (detection bias); D=incomplete outcome data (attrition bias); E=selective reporting (reporting bias)

# Steps in a systematic review

1. formulate the question
2. decide eligibility criteria
3. develop a protocol
4. search for studies
5. apply eligibility criteria to select relevant studies
6. collect data from the studies
7. assess studies for risk of bias/methodological quality
8. summarize the studies
9. synthesize findings of the studies, e.g. meta-analysis
10. interpret results and draw conclusions

### Intervention 3-6 years



Test for heterogeneity:  $\tau^2=0.03$ ,  $\chi^2=21.68$ ,  $df=12$ ,  $P=0.04$ ,  $I^2=45\%$

Test for overall effect:  $z=6.32$ ,  $P<0.001$

# Key ideas of meta-analysis

- Characteristics of patients likely to vary across studies
- Patients should only be compared with others in the same study
- A simple approach is to calculate a weighted average of treatment effects from each study
- We will continue with risk ratios for this example
  - Methods are similar for other measures of treatment effect

# Fixed-effect inverse-variance-weighted average (1)

- Require from each of  $k$  studies:
  - estimate of treatment effect,  $y_i$
  - estimate of variance of estimate,  $v_i$
  - (When using ratio measures, natural log of the ratio is used)
- Combine the estimates using a weighted average
- Take weight = inverse variance:
$$w_i = 1 / v_i$$
- It is intuitively sensible to give more weight to the bigger studies

# Application to risk ratios

- From each trial:

- Log risk ratio

$$y_i = \ln RR_i = \ln \frac{r_{Ai}/n_{Ai}}{r_{Bi}/n_{Bi}}$$

- Variance of log risk ratio  $\approx$

$$v_i = \frac{1}{r_{Ai}} - \frac{1}{n_{Ai}} + \frac{1}{r_{Bi}} - \frac{1}{n_{Bi}}$$

- Weight

$$w_i = \frac{1}{v_i}$$

	Event	No event	Total
Treatment A	$r_{Ai}$	$f_{Ai}$	$n_{Ai}$
Treatment B	$r_{Bi}$	$f_{Bi}$	$n_{Bi}$
.	.	.	.

# Application to Roumen 2008

	Event	No event	Total
Treatment A	11	63	74
Treatment B	19	54	73

- From each trial:

– Log risk ratio  $y_i = \ln RR_i = \ln \frac{11/74}{19/73} = \ln 0.571 = -0.56$

– Variance of log risk ratio  $\approx v_i = \frac{1}{11} - \frac{1}{74} + \frac{1}{19} - \frac{1}{73} = 0.116$

– Weight  $w_i = \frac{1}{0.116} = 8.6$

# Fixed-effect inverse-variance-weighted average

- Summary estimate  $\mu$

$$\mu = \frac{\sum w_j y_j}{\sum w_j}$$

- A standard error is:  $SE(\mu) = \sqrt{1/\sum w_j}$
- 95% confidence interval for the summary estimate:

$$\mu - 1.96\sqrt{1/\sum w_j} \text{ to } \mu + 1.96\sqrt{1/\sum w_j}$$

$y_i$  can be anything! logOR, logRR, RD, logHR, mean difference, standardised mean difference etc.

# The main assumption a fixed-effect meta-analysis is that:

The 'true' treatment effect is the same (fixed) in all studies

The observed treatment effect is the same (fixed) in all studies

The 'true' treatment effects in studies form a distribution with a mean and a variance

The observed treatment effects are unrelated

# Fixed-effect Mantel-Haenszel method

Exactly the same idea as in the inverse variance....

$$\mu = \frac{\sum w_i y_i}{\sum w_i}$$

$$w_i = \frac{bc}{a+b+c+d} \quad \text{for OR}$$

..but with different weights  $w_i = \frac{c(a+b)}{a+b+c+d}$  for RR

$$w_i = \frac{(a+b)(c+d)}{a+b+c+d} \quad \text{for RD}$$

SE( $\mu$ ) with Mantel-Haenszel a bit more complicated....

# Fixed-effect Mantel-Haenszel method

- Is more robust for small sample sizes, and rare outcomes if a fixed effect assumption is appropriate

# Fixed-effect inverse-variance

Study	r1	n1	r2	n2
Kosaka 2005	3	102	32	356
Eriksson 1991	17	181	16	79
Tuomilehto 2001	27	265	59	257

in R: `poolRR<-metabin(r1,n1,r2,n2,studlab=Study,sm="RR",`

`summary(poolRR)`

Mantel-Haenszel

	RR	95%-CI	z	p-value
Fixed effect model	0.6036	[0.5496; 0.6631]	-10.54	< 0.0001
Random effects model	0.6264	[0.5425; 0.7234]	-6.37	< 0.0001

Quantifying heterogeneity:

$\tau^2 = 0.0244$ ;  $H = 1.34$  [1.00; 1.85];  $I^2 = 43.9\%$  [0.0%; 70.7%]

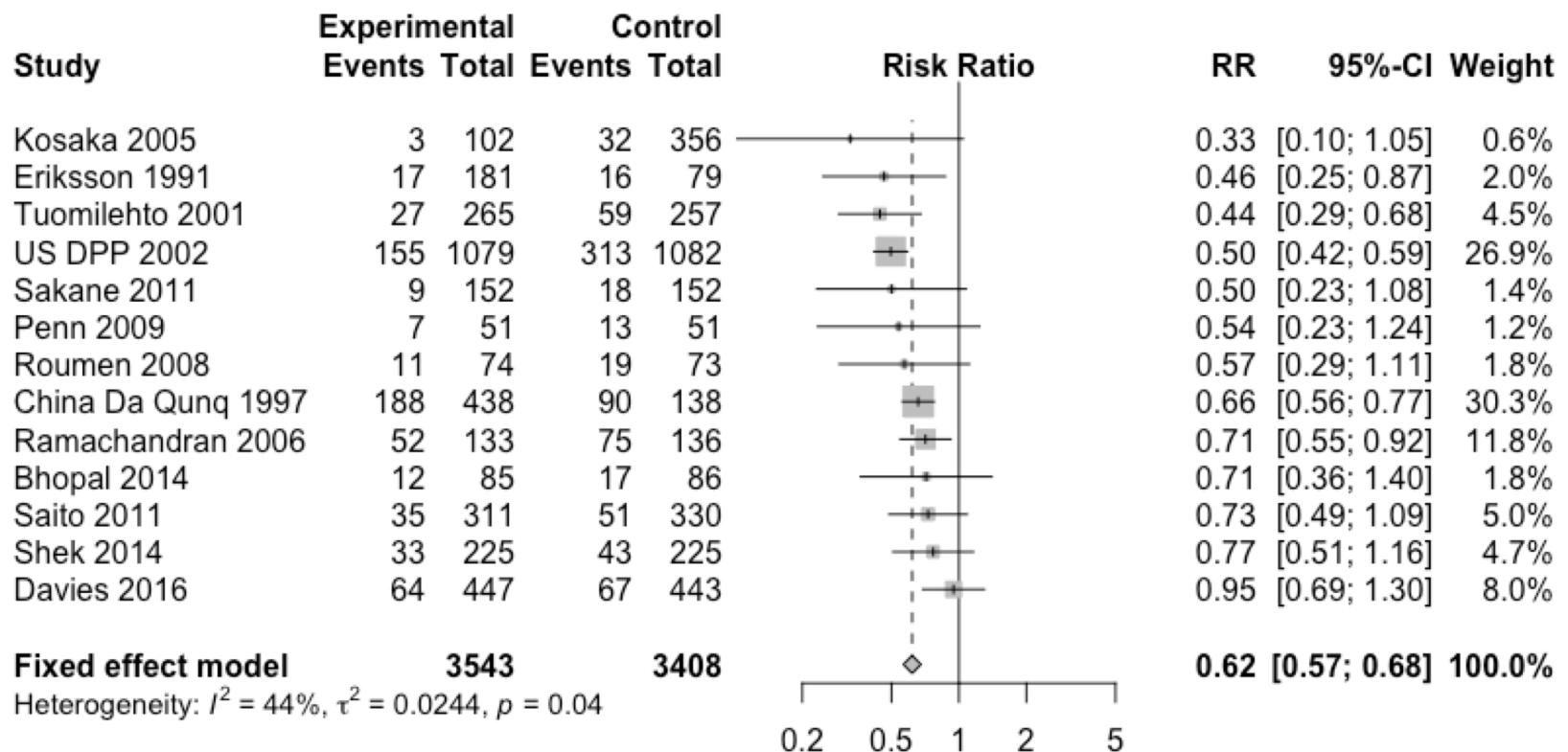
Test of heterogeneity:

Q	d.f.	p-value
21.39	12	0.0449

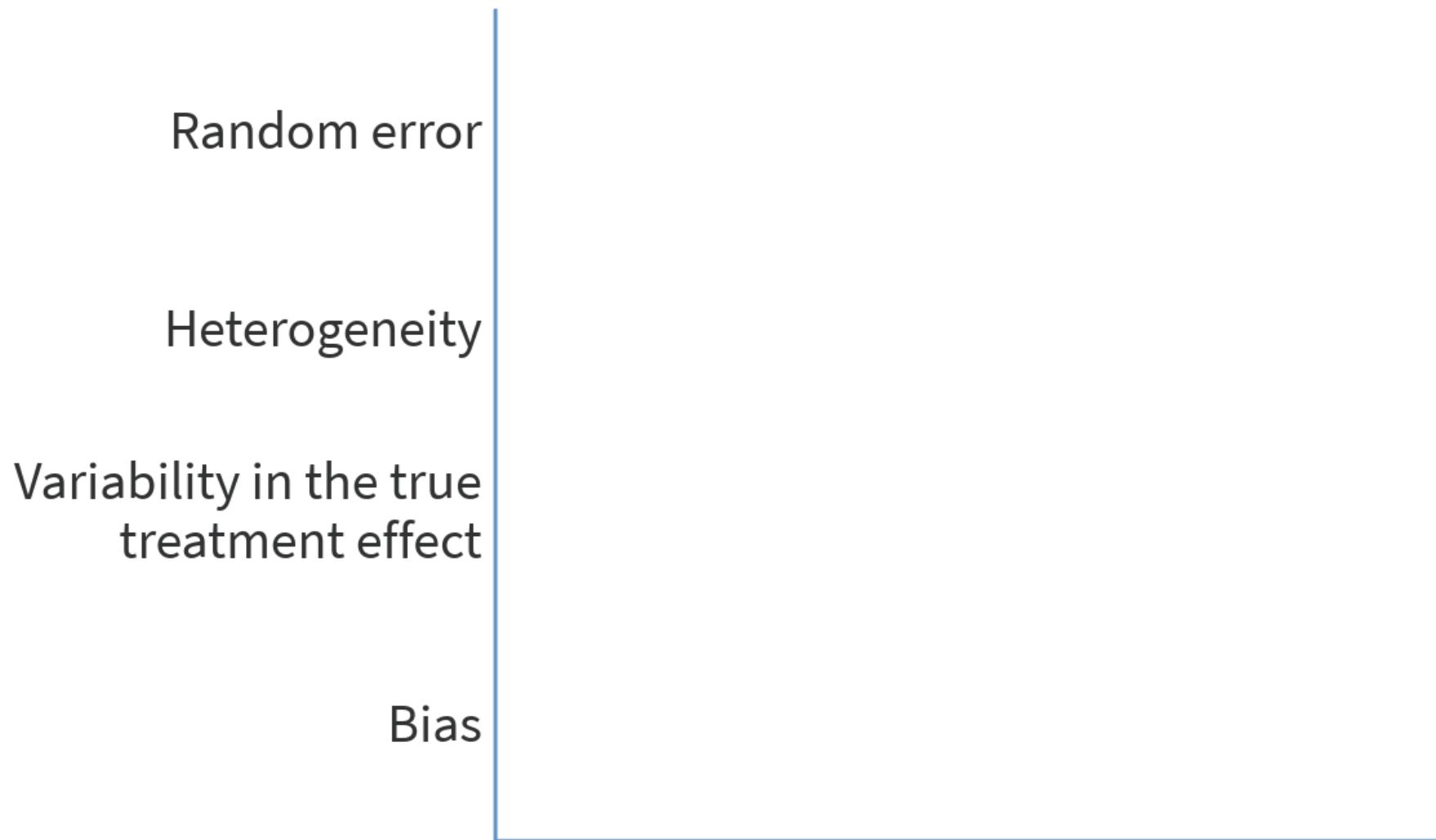
# Fixed-effect inverse-variance

Study	r1	n1	r2	n2
Kosaka 2005	3	102	32	356
Eriksson 1991	17	181	16	79
Tuomilehto 2001	27	265	59	257

in R:  
forest(poolRR)

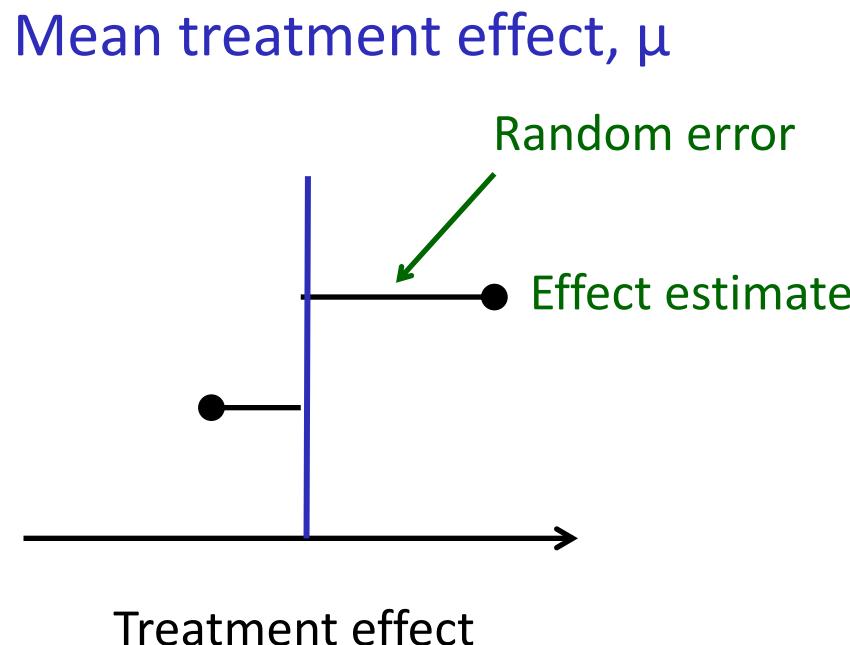


The observed treatment effects in the studies are different. The fixed effect model assumes that this is because of



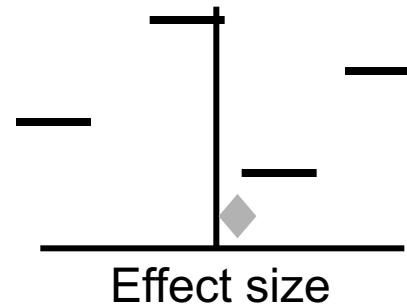
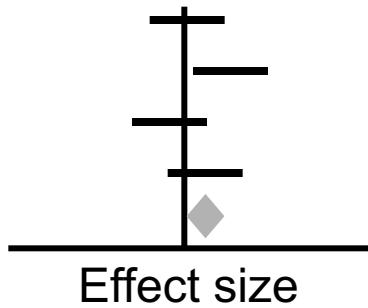
# Interpretation of fixed-effect meta-analysis results

- assume all studies share an identical true treatment effect
- estimate this single treatment effect



# Interpretation of fixed-effect meta-analysis results

- The confidence interval for the summary odds ratio reflects *within-study errors only*
- Variation across studies (heterogeneity) is ignored
- So the following yield exactly the same result



- Many of us feel uncomfortable about this

# What is the main assumption in a random-effects meta-analysis

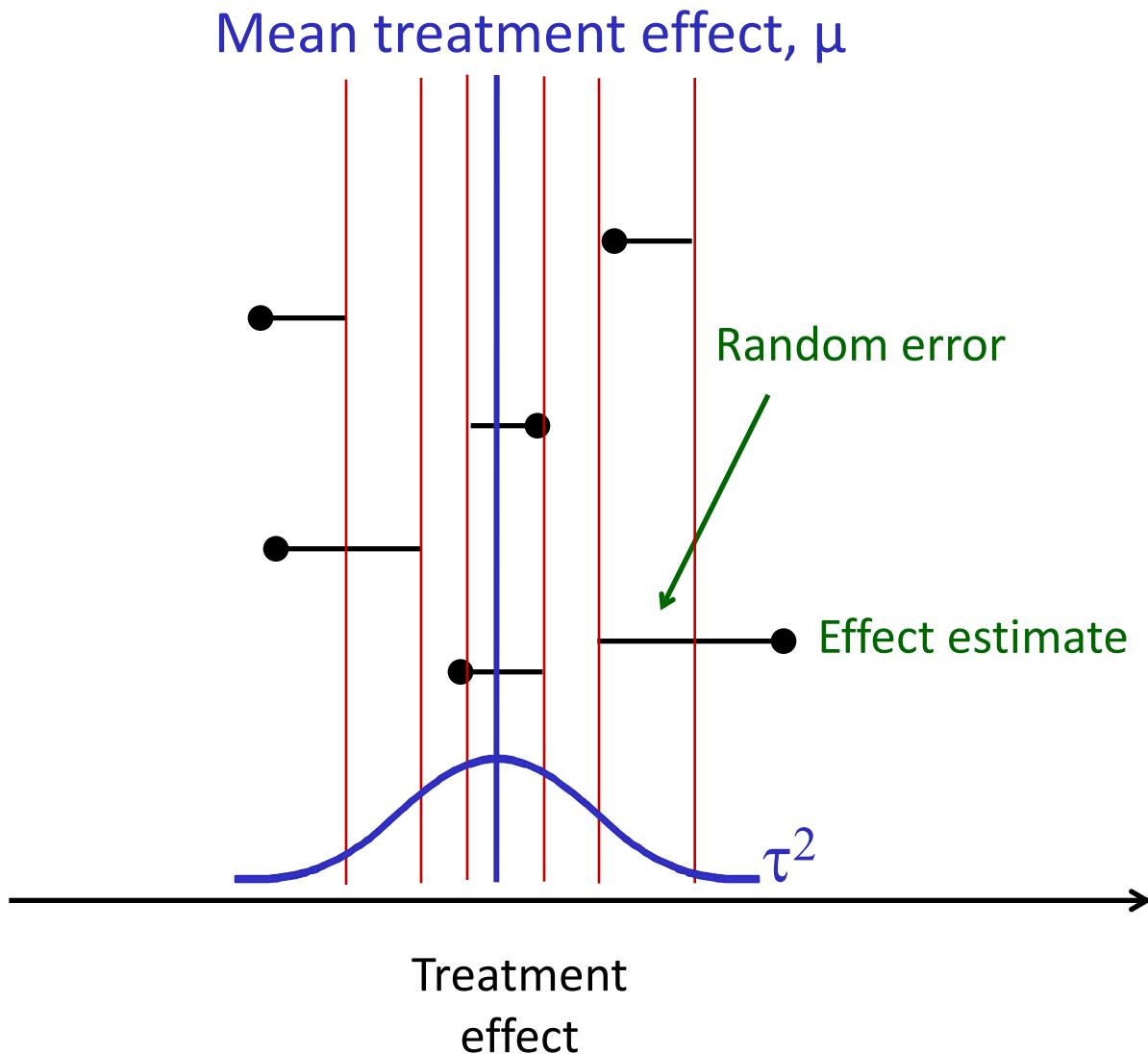
The 'true' treatment effects in studies are random and unrelated

The observed treatment effects in studies are random and unrelated

The 'true' treatment effects are related via a distribution with a mean and a variance

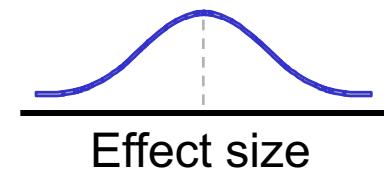
There is random error in the observed treatment effects

# Random-effects meta-analysis



# Random-effects meta-analysis

- We suppose the *true* treatment effect in each study is randomly, normally distributed across studies
  - with variance  $\tau^2$  (“tau-squared”)
- Estimate the between-study variance  $\tau^2$ , and use this to modify the weights used to calculate the summary estimate
- The most common estimate of  $\tau^2$  is called the DerSimonian and Laird estimate, or method of moments estimate
  - but other (better) estimators are available



## Random-effects meta-analysis (3)

Random-effects estimate:  $\mu = \frac{\sum w_i^* y_i}{\sum w_i^*}$

where  $w_i^* = \frac{1}{v_i + \tau^2}$

A standard error is:  $SE(\mu) = \sqrt{1/\sum w_i^*}$

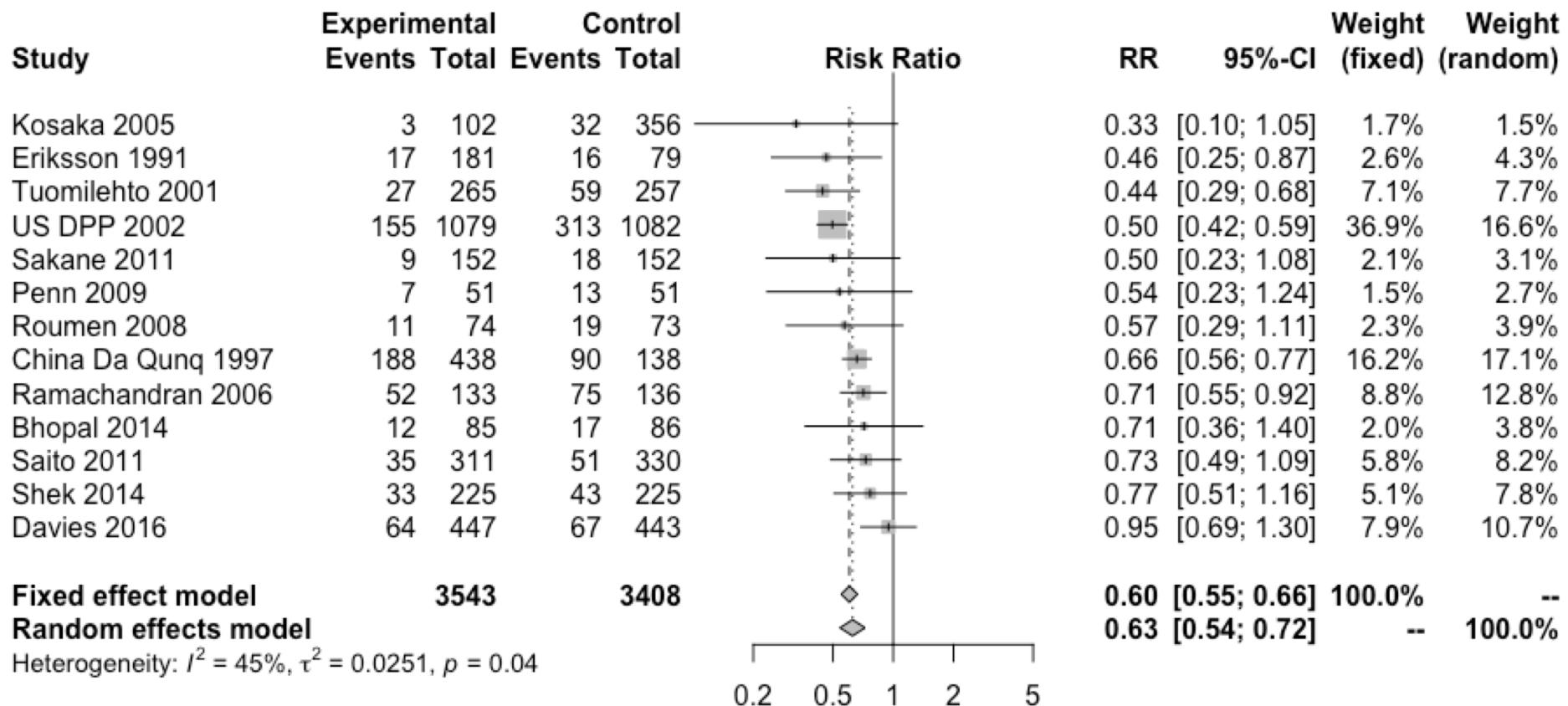
95% confidence interval for the summary estimate:

$$\mu - 1.96 \sqrt{1/\sum w_i^*} \text{ to } \mu + 1.96 \sqrt{1/\sum w_i^*}$$

We estimate both random and fixed effects meta-analyses. The mean treatment effect is the same. Heterogeneity variance is relatively large. Then, for the confidence intervals of the summary effect we expect that that

- The 95% CI from the FE is wider than the 95% CI from the RE      A
- The 95% CI from the RE is wider than the 95% CI from the FE      B
- The two 95% CI do not overlap      C
- The two 95% CIs have 95% overlap      D

# Example (continued)



# Identifying heterogeneity : test

- To test the null hypothesis that the true treatment effect is the same in all studies we can calculate a *heterogeneity statistic*:

$$Q = \sum w_i(y_i - \mu)^2$$

- To calculate a P value,  $Q$  is compared with the  $\chi^2$  distribution on  $(k - 1)$  degrees of freedom ( $k$  is no. of studies).
- The greater the average weighted squared distance between the individual study log risk ratio  $y_i$  and the summary log risk ratio  $\mu$ , the more evidence against the null hypothesis that the true treatment effect is the same in all studies.

# Identifying heterogeneity: $I^2$

- An alternative measure to quantify inconsistency
  - based on  $\chi^2$  statistic,  $Q$ , and its degrees of freedom ( $= k - 1$ ).

$$I^2 = \frac{Q - \text{d.f.}}{Q} \times 100\%$$

- $I^2$  can be interpreted as the proportion of total variability explained by heterogeneity, rather than chance
- If  $I^2$  is computed to be  $< 0$ , it is set to 0%

# Identifying heterogeneity: $\tau^2$

- The between-studies variance,  $\tau^2$  is estimated as part of the random-effects meta-analysis
- It provides a useful measure of the true extent of heterogeneity across studies
- Methods to estimate  $\tau^2$ 
  - **DerSimonian and Laird** estimator is default; `method.tau="DL"`
  - **Paule-Mandel**; `method.tau="PM"` (Paule and Mandel, 1982)
  - **Restricted maximum-likelihood**; `method.tau="REML"` (Viechtbauer, 2005)
- Simulations and empirical analyses suggest that for both dichotomous and continuous data PM and for continuous data REML are better alternatives (Veroniki et al., Res Synth Meth 2015)

# Example

## Mantel-Haenszel method

Number of studies combined:  $k = 13$

	RR	95%-CI	z	p-value
Fixed effect model	0.6036 [0.5496; 0.6631]	-10.54 < 0.0001		
Random effects model	0.6263 [0.5418; 0.7241]	-6.32 < 0.0001		

Quantifying heterogeneity:

$\tau^2 = 0.0251$ ;  $H = 1.34$  [1.00; 1.86];  $I^2 = 44.7\%$  [0.0%; 71.1%]

Test of heterogeneity:

Q	d.f.	p-value
21.68	12	0.0412

Details on meta-analytical method:

- Mantel-Haenszel method
- DerSimonian-Laird estimator for  $\tau^2$

Some evidence that the effect of intervention differs between studies  
29

# Example: Compare estimators of $\tau^2$

Inverse variance RE method

**DerSimonian and Laird**

	RR	95%-CI	z	p-value
Random effects model	0.6264	[0.5425; 0.7234]	-6.37	< 0.0001

Quantifying heterogeneity:

$\tau^2 = 0.0244$ ;  $H = 1.34$  [1.00; 1.85];  $I^2 = 43.9\%$  [0.0%; 70.7%]

Test of heterogeneity:

Q	d.f.	p-value
21.39	12	0.0449

**Paule-Mandel**

	RR	95%-CI	z	p-value
Random effects model	0.6269	[0.5489; 0.7160]	-6.89	< 0.0001

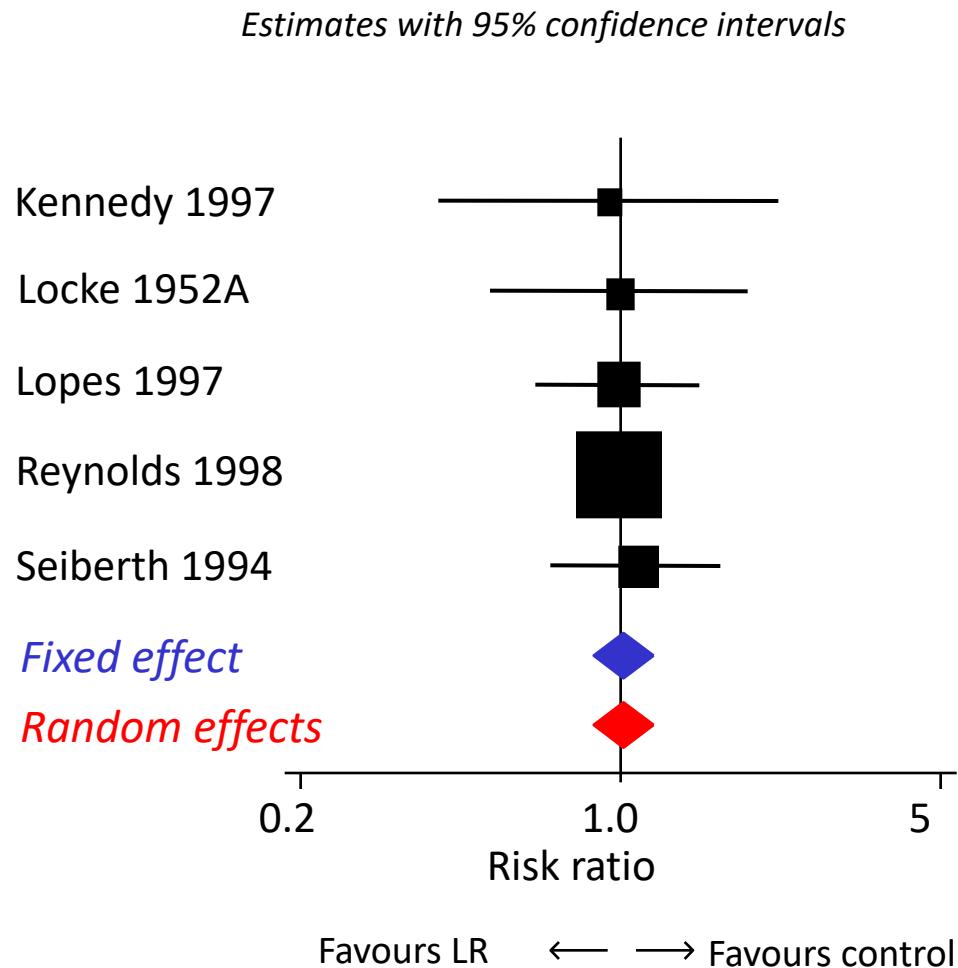
Quantifying heterogeneity:

$\tau^2 = 0.0174$ ;  $H = 1.34$  [1.00; 1.85];  $I^2 = 43.9\%$  [0.0%; 70.7%]

Test of heterogeneity:

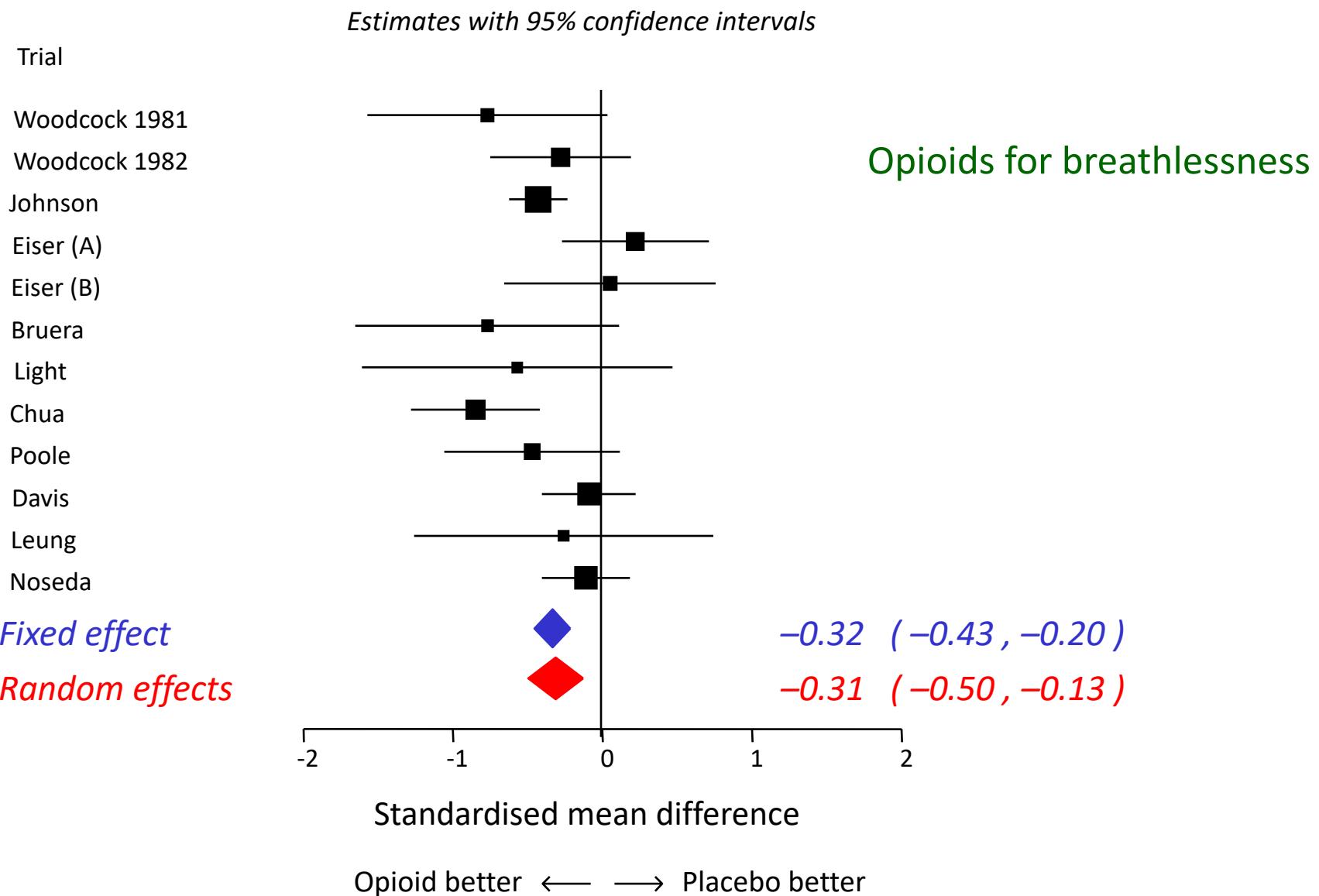
Q	d.f.	p-value
21.39	12	0.0449

# Fixed versus random effects: Identical results

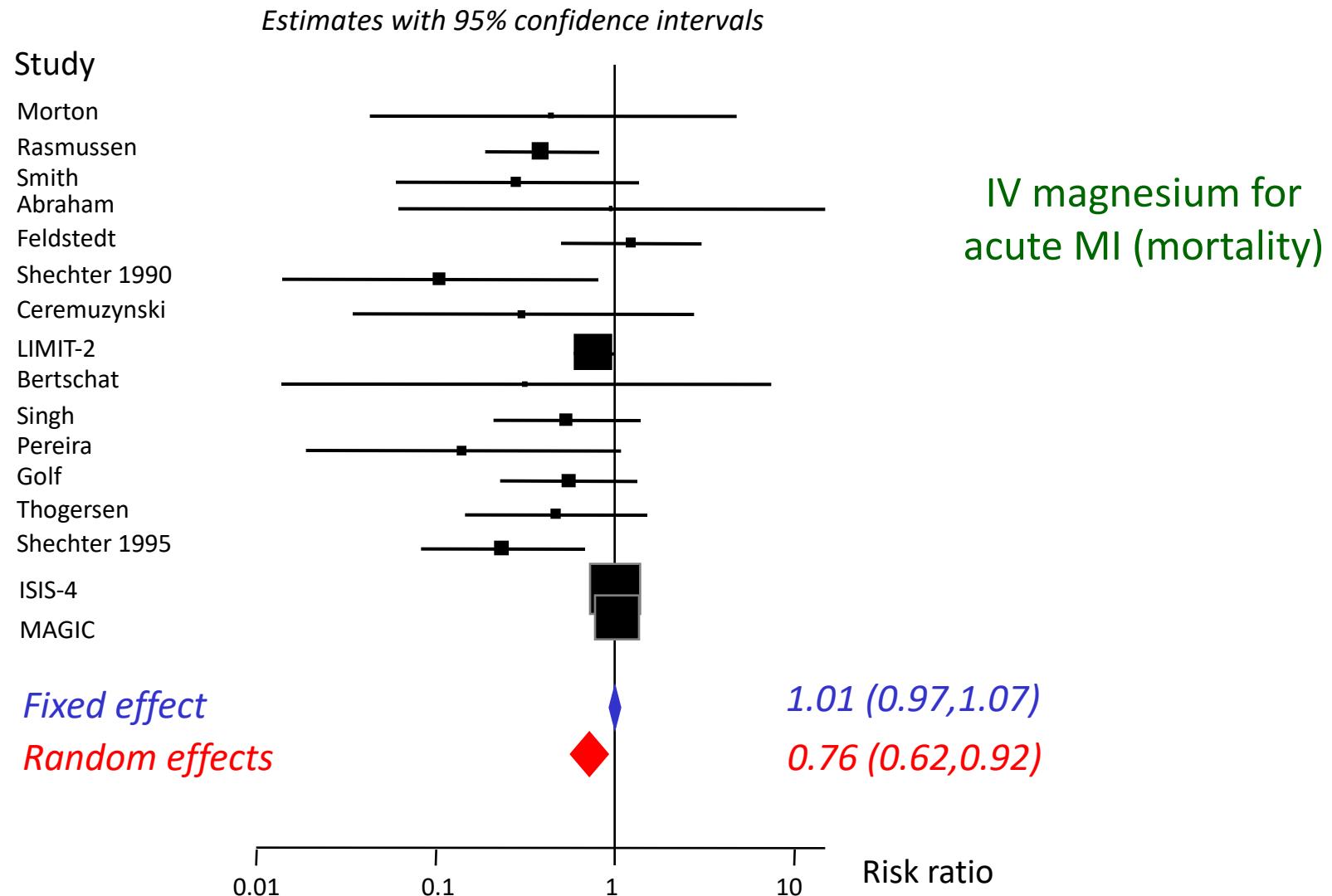


Early light reduction for preventing retinopathy of prematurity

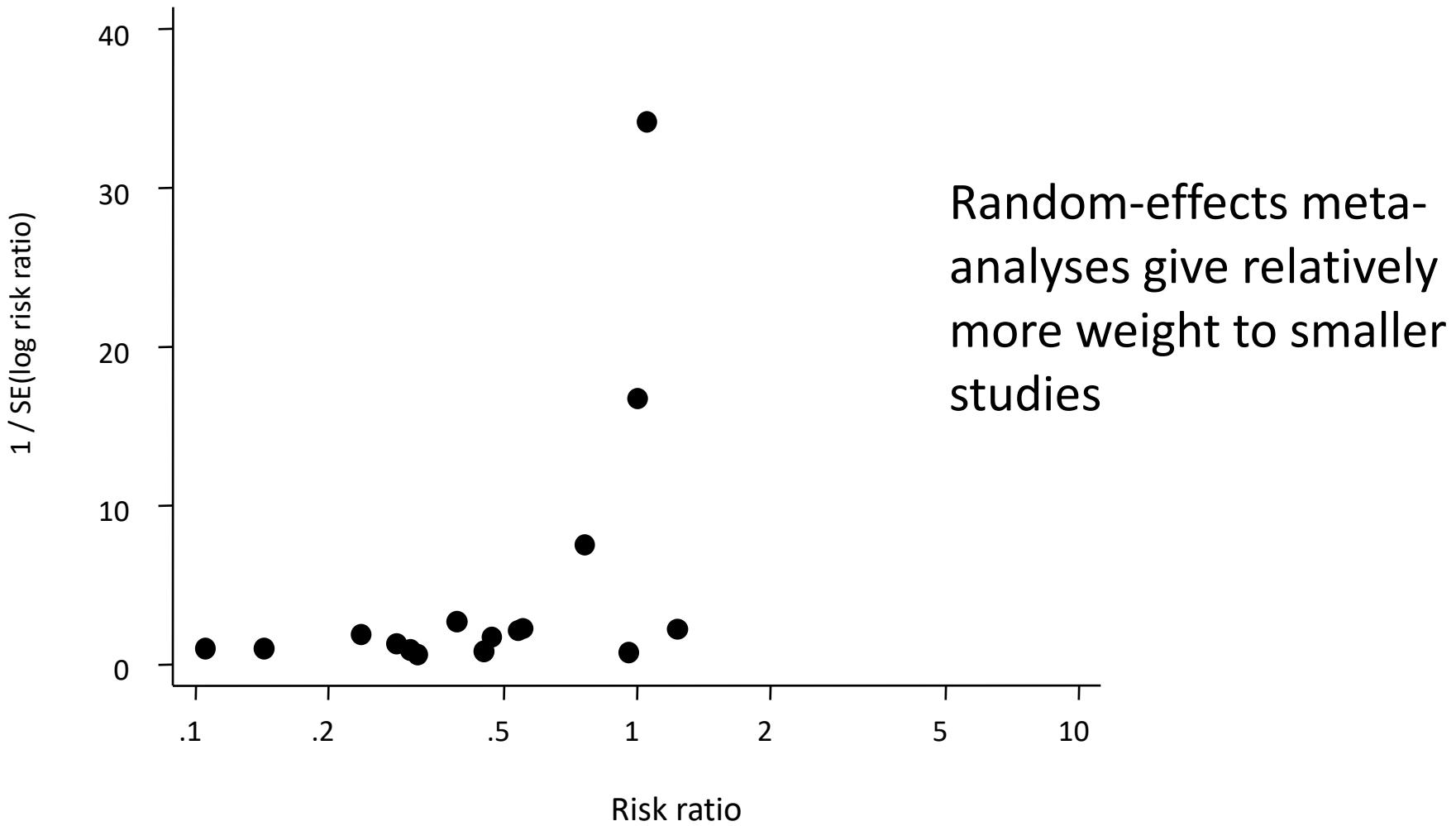
# Fixed versus random effects: Slightly different results



# Fixed versus random effects: Very different results



# Funnel plot: asymmetrical



# Random-effects meta-analysis

Random-effects estimate:  $\mu = \frac{\sum y_i}{\sum w_i^* y_i}$

where  $w_i^* = \frac{1}{v_i + \tau^2}$

A standard error is:  $SE(\mu) = \sqrt{1/\sum w_i}$

# Random-effects meta-analysis

## Hartung-Knapp adjustment

A standard error of the weighted mean  $\mu$  is

$$SE(\mu) = \sqrt{\frac{\sum \frac{w_i^*}{\sum w_j^*} (y_i - \mu)^2}{k-1}}$$

A 95% CI for  $\mu$  is

$$\mu - t_{k-1, 0.025} SE(\mu) \quad \text{to} \quad \mu + t_{k-1, 0.975} SE(\mu)$$

# Example:

## Compare 95% Confidence Intervals for $\mu$

### Inverse variance method

**RE model with Paule-Mandel estimator for  $\tau^2$**

```
poolRR<-metabin(r1,n1,r2,n2,studlab=Study,sm="RR", method.tau="PM",method="I")
```

	RR	95%-CI	z	p-value
Random effects model	0.6269	[0.5489; 0.7160]	-6.89	< 0.0001

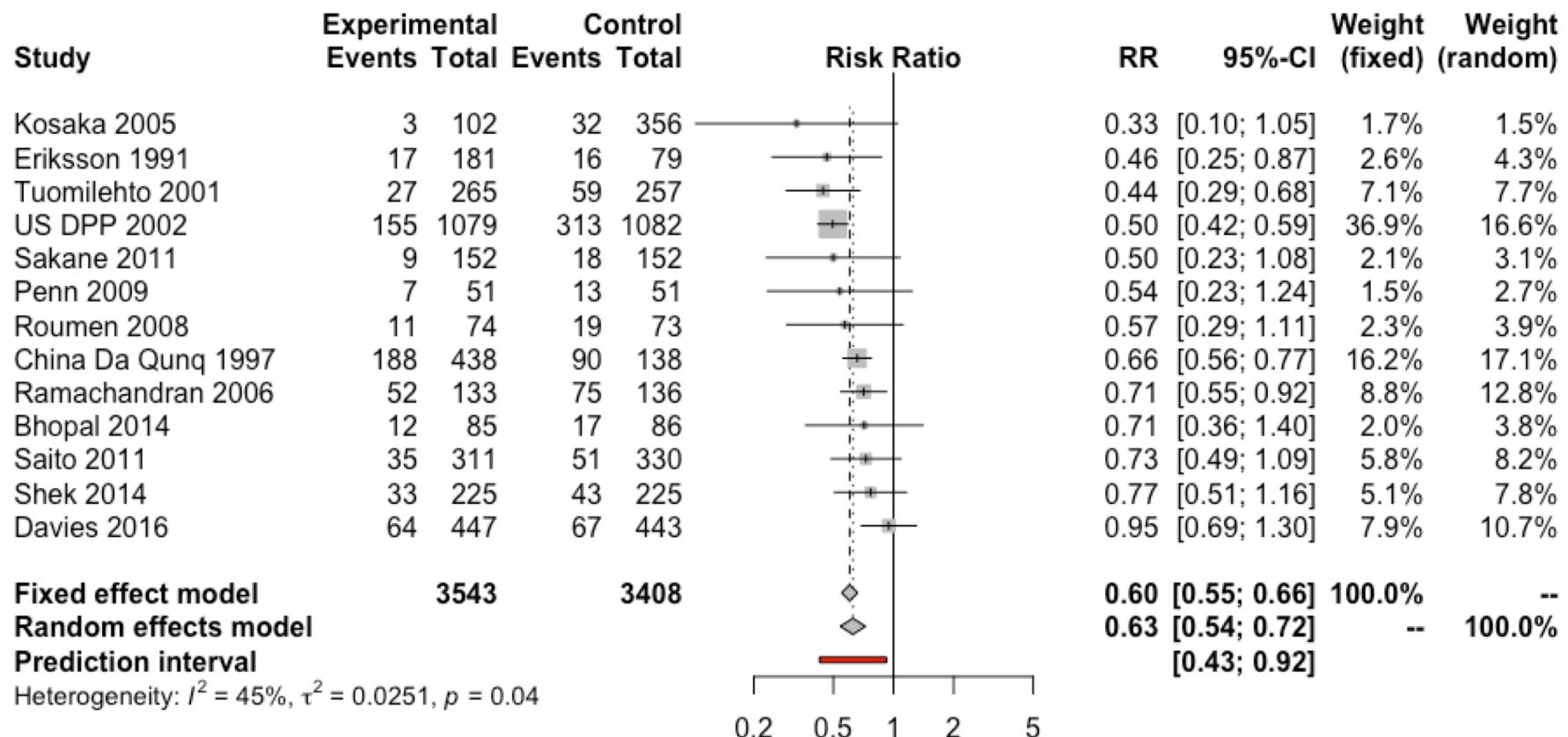
**RE model with Paule-Mandel estimator for  $\tau^2$   
and Hartung-Knapp adjustment**

```
update(poolRR,hakn=T)
```

	RR	95%-CI	t	p-value
Random effects model	0.6269	[0.5408; 0.7267]	-6.89	< 0.0001

# Add random-effects prediction interval

**forest(poolRR, prediction=T)**



# Prediction intervals

- Prediction intervals portray the actual heterogeneity across studies
- An interval within which the true effect size in a similar study (from the same distribution) is predicted to lie

$$\mu - t_{k-1, 0.025} \sqrt{SE(\mu)^2 + \tau^2} \text{ to } \mu + t_{k-1, 0.975} \sqrt{SE(\mu)^2 + \tau^2}$$

# This is a generic approach

- Suppose we have an estimate of some quantity,  $y_i$ , and we know its variance,  $v_i$
- Then we can perform a meta-analysis as a weighted average

FE:

$$\hat{\theta} = \frac{\sum w_i y_i}{\sum w_i}$$

RE:

$$\hat{\mu} = \frac{\sum w_i^* y_i}{\sum w_i^*}$$

$$SE(\hat{\theta}) = \sqrt{\frac{1}{\sum w_i}}$$

$$SE(\hat{\mu}) = \sqrt{\frac{1}{\sum w_i^*}}$$

- For example, for binary data,  $y_i$  could be a log odds ratio or a risk difference
  - we (almost) always work on the log scale for ratio measures
- Note: these methods ignore uncertainty in  $\tau^2$

# Review: basic principles of meta-analysis (1)

- Start with a clear question
- Compare like with like
  - participants in one study are not directly compared with those in another
  - each study is analysed separately
  - summary statistics are combined to give the meta-analysis
- Weight studies according to the information they provide
  - usually by precision
  - gives more weight to larger studies...
  - ... so that larger studies have more influence on the summary estimate

# References

- Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-analysis*. John Wiley & Sons, 2009.
- Collins R, Yusuf S, Peto R. Overview of randomised trials of diuretics in pregnancy. *Br Med J (Clin Res Ed)* 1985; **290**: 17–23.
- Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M, Davey Smith G, Altman DG (editors). *Systematic Reviews in Health Care: Meta-analysis in Context* (2nd edition). London (UK): BMJ Publication Group, 2001.
- DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986; **7**: 177-188.
- Hartung J, Knapp G. A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Stat Med* 2001; **20**: 3875-89.
- Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society Series A* 2009; **172**: 137-159.
- Lawlor DA, Hopker SW. The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. *BMJ* 2001; **322**: 763-7.
- Whitehead A. (2002) *Meta-analysis of Controlled Clinical Trials*. Chichester: Wiley.
- Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomised clinical trials. *Stat Med* 1991; **10**: 1665-77.