Improving coverage probabilities of confidence intervals in random effects meta-analysis with publication bias

Masayuki Henmi
The Institute of Statistical Mathematics, Japan
John B. Copas
University of Warwick, UK

Introduction

- Meta-analysis: statistical analysis to strengthen some (statistical) evidence by combining results from several studies

  Clinical trials, Epidemiological studies, etc.

- Study heterogeneity: difference of treatment effect in each study

- DerSimonian-Laird method: a popular method to deal with study heterogeneity, but the confidence interval from this method has **poor coverage probabilities**, especially in the presence of **publication bias**

In this talk, we propose a new confidence interval which improves such poor coverage probabilities, taking into account the effect of publication bias.

Fixed effects model

Fixed effects $\theta_1 = \theta_2 = \cdots = \theta_n = \theta$

θ_i : treatment effect of the ith study

Estimate of θ from the ith study $y_i \sim N\left(\theta, s_i^2\right)\ \left(i = 1, 2, \ldots n\right)$

Overall estimate $\hat{\theta}_F = \dfrac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}\ \left(w_i = 1/s_i^2\right)$

100(1- α)% confidence interval

$$\left[\hat{\theta}_F - z_{\alpha/2}\left(\sum_{i=1}^{n} w_i\right)^{-1/2}, \hat{\theta}_F + z_{\alpha/2}\left(\sum_{i=1}^{n} w_i\right)^{-1/2}\right]$$

($z_{\alpha/2}$ is the $\alpha/2$ upper quantile of $N(0,1)$)

Random effects model
—DerSimonian-Laird method—

Usual method to deal with study heterogeneity $i.e.\ \theta_j \neq \theta_k \ (j \neq k)$

Random effects $\theta_i \sim N(\theta, \tau^2) \longrightarrow y_i \sim N(\theta, s_i^2 + \tau^2)$

Overall estimates $\hat{\theta}_R = \dfrac{\sum_{i=1}^n \hat{w}_i y_i}{\sum_{i=1}^n \hat{w}_i} \quad \left(\hat{w}_i = 1\Big/\left(s_i^2 + \hat{\tau}_{\mathrm{DL}}^2\right)\right)$

↑ DerSimonian-Laird estimator

By identifying $\hat{\tau}_{\mathrm{DL}}^2$ with τ^2, $T_{\mathrm{DL}} = \dfrac{\hat{\theta}_R - \theta}{\sqrt{\hat{V}_{\mathrm{DL}}}} \sim N(0,1) \quad \left(\hat{V}_{\mathrm{DL}} = \dfrac{1}{\sum_{i=1}^n \hat{w}_i}\right)$

$100(1-\alpha)$ %

Confidence interval $I_{\mathrm{DL}} = \left[\hat{\theta}_R - z_{\alpha/2}\sqrt{\hat{V}_{\mathrm{DL}}},\ \hat{\theta}_R + z_{\alpha/2}\sqrt{\hat{V}_{\mathrm{DL}}}\right]$

Problems of DerSimonian-Laird method

Confidence interval $I_{\mathrm{DL}} = \left[\hat{\theta}_R - z_{\alpha/2}\sqrt{\hat{V}_{\mathrm{DL}}},\ \hat{\theta}_R + z_{\alpha/2}\sqrt{\hat{V}_{\mathrm{DL}}} \right]$

1) The coverage probability tends to be below the nominal level $(1-\alpha)$, if we take into account the effect of estimating τ^2 (Brockwell and Gordon, 2001)

<u>Alternative confidence intervals</u>

- Likelihood ratio (Hardy and Thompson, 1996)
- Biggerstaff and Tweedie (1997)
- Sidik and Jonkman (2002)

⋮

Problems of DerSimonian-Laird method

Confidence interval $I_{\mathrm{DL}} = \left[\hat{\theta}_R - z_{\alpha/2}\sqrt{\hat{V}_{\mathrm{DL}}},\ \hat{\theta}_R + z_{\alpha/2}\sqrt{\hat{V}_{\mathrm{DL}}} \right]$

2) If there is publication bias, the coverage probability decreases more rapidly.

Publication bias: small studies with large standard errors s_i are less likely to
(typical case) be published than large studies with small standard errors.

The random effects estimate $\hat{\theta}_R = \dfrac{\sum_{i=1}^{n} \hat{w}_i y_i}{\sum_{i=1}^{n} \hat{w}_i} \left(\hat{w}_i = \dfrac{1}{s_i^2 + \hat{\tau}_{\mathrm{DL}}^2} \right)$ is more sensitive

to publication bias than the fixed effects estimate $\hat{\theta}_F = \dfrac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i} \left(w_i = \dfrac{1}{s_i^2} \right)$,

because $s_i > s_j \quad \Longrightarrow \quad \dfrac{\hat{w}_i}{\hat{w}_j} = \dfrac{s_j^2 + \hat{\tau}_{\mathrm{DL}}^2}{s_i^2 + \hat{\tau}_{\mathrm{DL}}^2} > \dfrac{s_j^2}{s_i^2} = \dfrac{w_i}{w_j}$

relative weights of smaller studies

Confidence interval
based on the fixed effects estimate

Under $y_i \sim N\left(\theta, s_i^2 + \tau^2\right)$, $V_{\mathrm{HC}} = \mathrm{Var}\left(\hat{\theta}_F\right) = \dfrac{\tau^2 \sum_{i=1}^n w_i^2 + \sum_{i=1}^n w_i}{\left(\sum_{i=1}^n w_i\right)^2}$

$$\Longrightarrow \quad T_{\mathrm{HC}} = \frac{\hat{\theta}_F - \theta}{\sqrt{\hat{V}_{\mathrm{HC}}}} \quad \left(\hat{V}_{\mathrm{HC}} = \frac{\hat{\tau}_{\mathrm{DL}}^2 \sum_{i=1}^n w_i^2 + \sum_{i=1}^n w_i}{\left(\sum_{i=1}^n w_i\right)^2} \right)$$

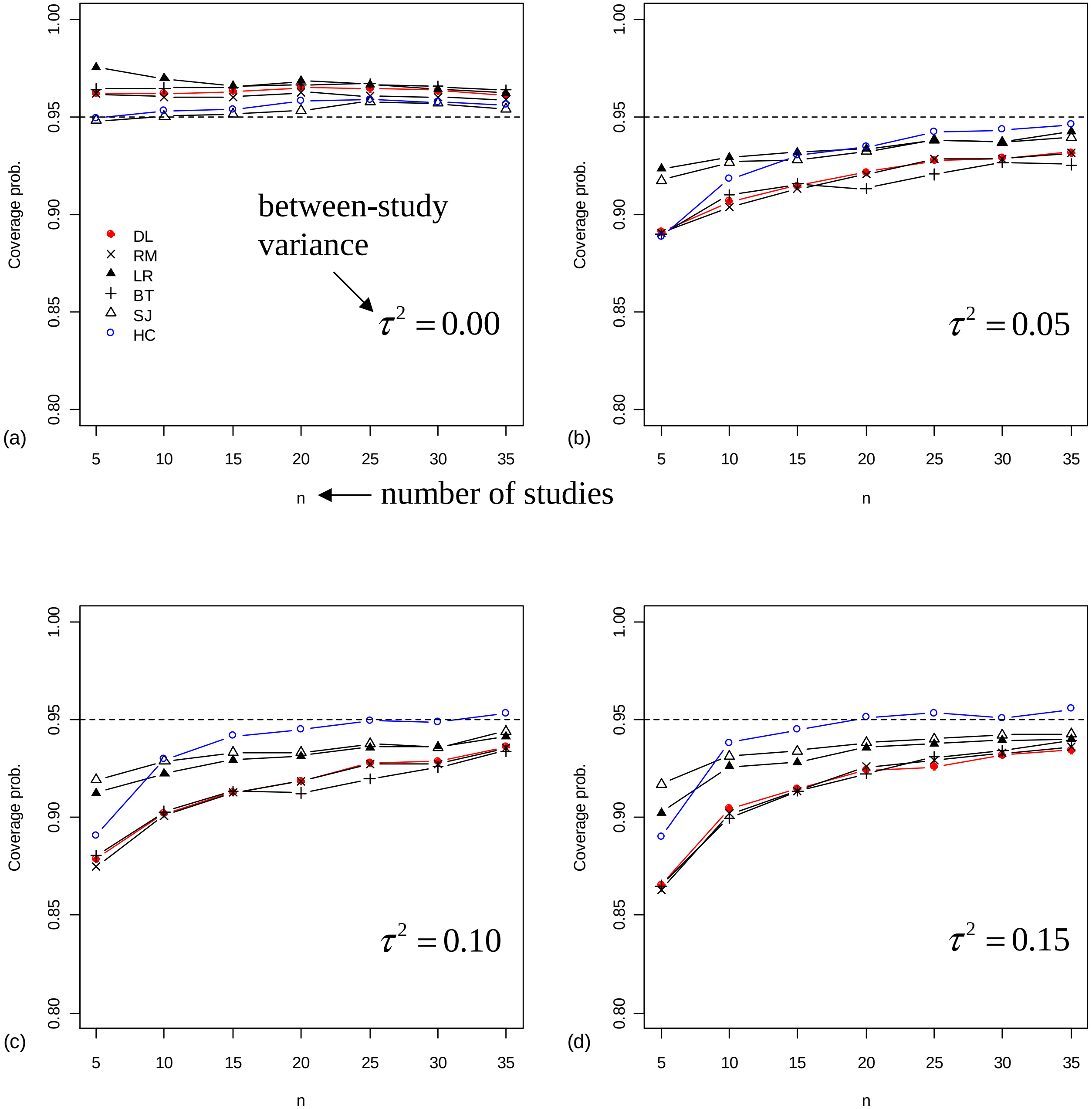
Calculate **the exact distribution** to obtain **more accurate approximation** than the approximation by normal distribution

$\Longrightarrow$ approximate 100(1- α)% confidence interval

$$I_{\mathrm{HC}} = \left[\hat{\theta}_F - u_{\alpha/2}\sqrt{\hat{V}_{\mathrm{HC}}},\ \hat{\theta}_F + u_{\alpha/2}\sqrt{\hat{V}_{\mathrm{HC}}} \right] \quad \text{（ } u_{\alpha/2} \text{ : } \alpha/2 \text{ upper quantile of } T_{\mathrm{HC}} \text{ ）}$$

calculated from the above distribution but it depends on τ^2,
which is replaced with $\hat{\tau}_{\mathrm{DL}}^2$ here.

Simulation (no publication bias)



Overall average effect

$$\theta = 0.5$$

Within-study variance s_i^2 is generated by the method of Brockwell and Gordon (2001)

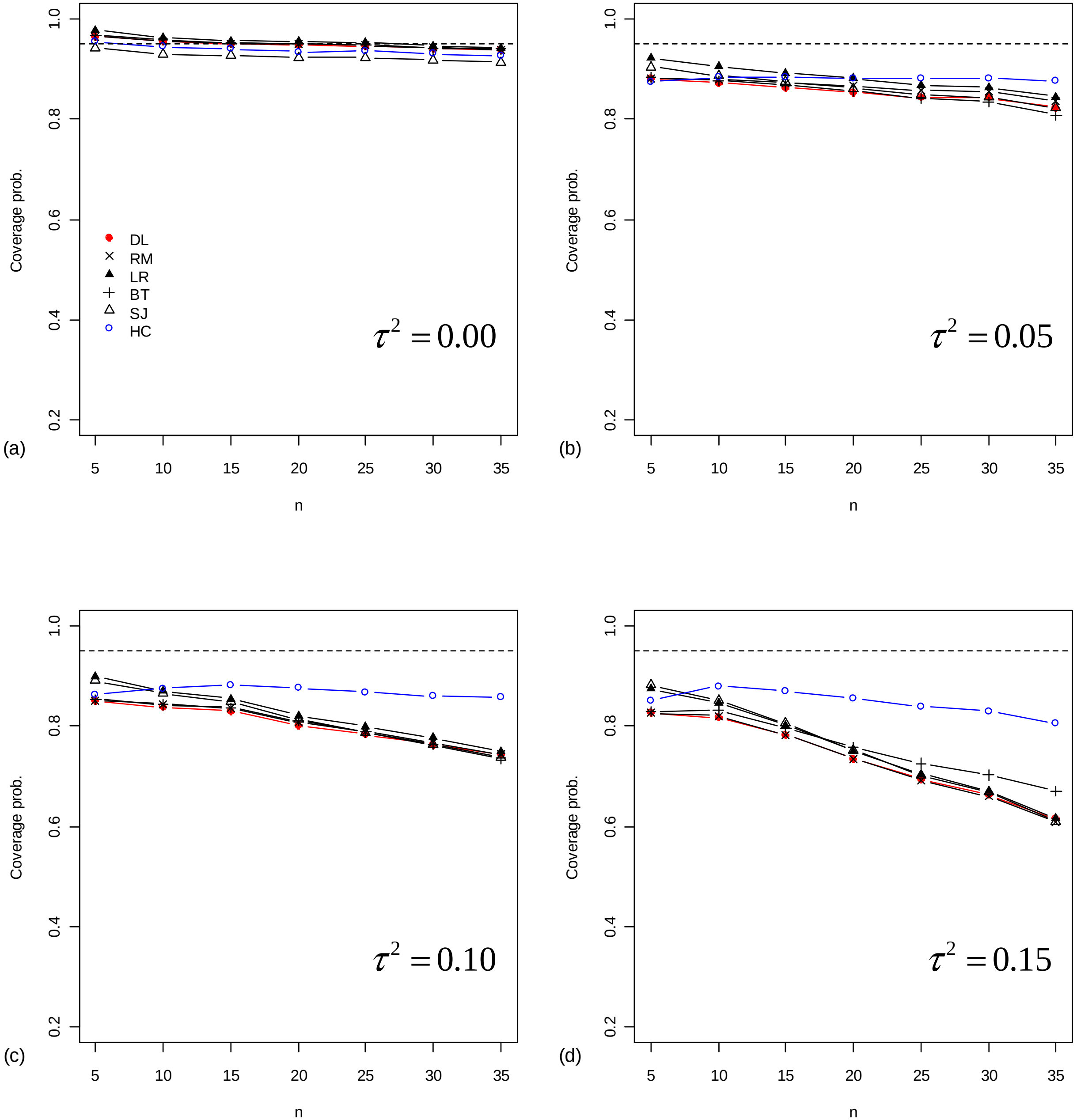
Observed effect size from each study

$$y_i \sim N\left(\theta, s_i^2 + \tau^2\right)$$

$$\left(i = 1, \mathrm{K}\ , n\right)$$

Simulation size 10000 times

Simulation (moderate publication bias)



Selection function

$$P(\text{selected} \mid y, s) = \exp\left[-\beta\left\{\Phi\left(-\frac{y}{s}\right)\right\}^{\gamma}\right]$$

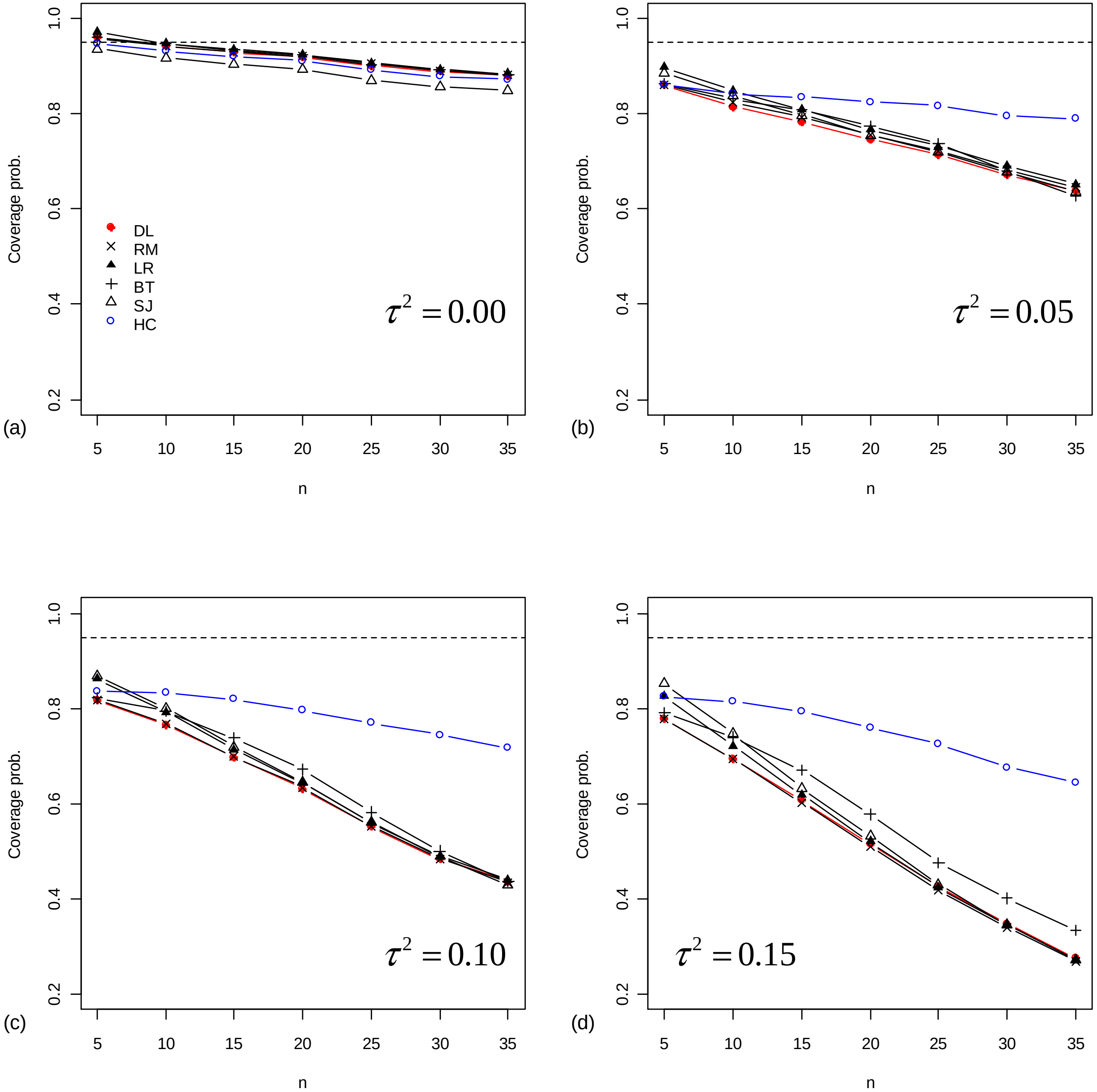
one-sided P-value

$\beta = 4, \gamma = 3$

⇒ moderate publication bias (average selection probability 87%)

(Begg and Mazumdar, 1994)

Simulation (strong publication bias)



Selection function

$$P(\text{selected} \mid y, s)$$

$$= \exp\left[-\beta \left\{ \Phi\left(-\frac{y}{s} \right) \right\}^{\gamma} \right]$$

one-sided P-value

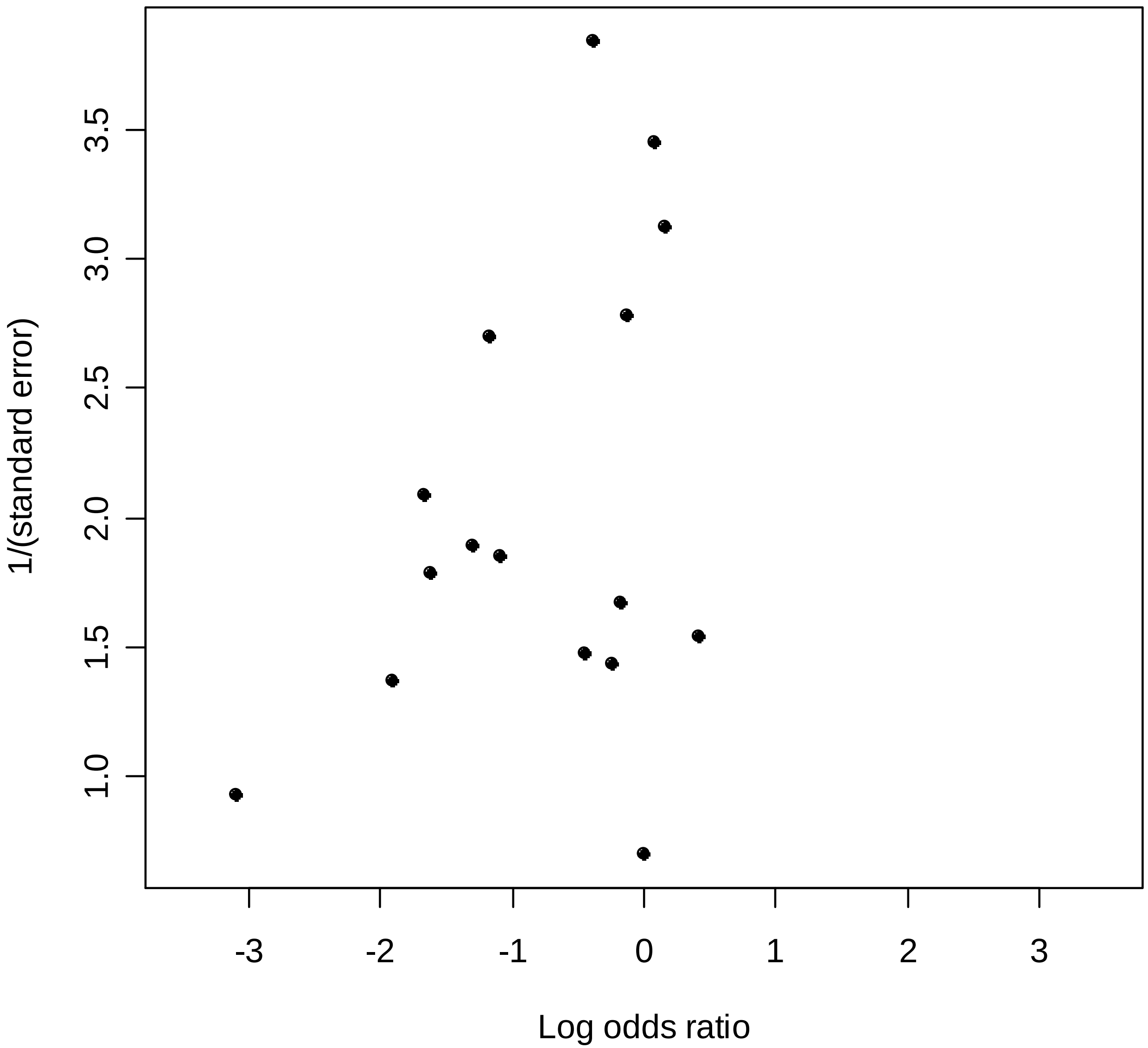
$\beta = 4,\ \gamma = 1.5$

⇒ strong publication bias

(average selection probability 76%)

(Begg and Mazumdar, 1994)

Clinical trials of wrist P6 acupoint stimulation in preventing postoperative nausea



Random effects estimate

$$\hat{\theta}_R = -0.68$$

Between-study variance

$$\hat{\tau}^2_{\text{DL}} = 0.33$$

Confidence intervals

DL $(-1.06, -0.29)$

RM $(-1.08, -0.29)$

BT $(-1.43, -0.18)$

SJ $(-1.12, -0.24)$

HC $(-1.00, \underline{-0.03})$

Summary

- DerSimonian-Laird confidence interval:
  lower coverage probability than the nominal level (95%)
  drastically lower if there is publication bias

- Fixed effects estimate:
  less sensitive to publication bias than the random effects estimate
  easier to calculate the exact distribution of the pivotal quantity

- Confidence interval based on the fixed effects estimate:
  higher coverage probability than the DL regardless of publication bias
  its difference becomes larger if publication bias exists

This confidence interval does not adjust for publication bias completely,
but gives more reliable information on the overall average effect
without making untestable assumption on the process of study selection.

References

Brockwell S.E. and Gordon I.R. (2001). A comparison of statistical methods for meta-analysis. *Statistics in Medicine* 20, 825-840.

DerSimonian R. and Laird N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials* 7, 177-188.

Hardy R.J. and Thompson S.G. (1996). A likelihood approach to meta-analysis with random effects. *Statistics in Medicine* 15, 619-629.

Biggerstaff B. J. and Tweedie R.L. (1997). Incorporating variability in esyimates of heterogeneity in the random effects model in meta-analysis. *Statistics in Medicine* 16, 753-768.

Sidik K. and Jonkman J.N. (2002). A simple confidence interval for meta-analysis. *Statistics in Medicine* 21, 3153-3159.

Begg C.B. and Mazumdar M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics* 50, 1088-1101.