MAY 14-16, 2007

LONDON, ENGLAND

The Seventh Annual International Campbell Collaboration Colloquium

Quality, Credibility, and Utility: The Relevance of Systematic Reviews

Symposium: Advances in Meta-analysis

Coordinator: Julio Sánchez Meca

Estimating an Overall Effect Size: Does the Between-Studies Variance Estimator have an Effect?

Fulgencio Marín Martínez and Julio Sánchez Meca

Universidad de Murcia

http://www.um.es/facpsi/metaanalysis

This research was supported by a Grant of the Ministerio de Educación y Ciencia and by FEDER Founds (Project nº: SEJ2004-0727B).

1. Three main objectives in meta-analysis

1. To derive an average effect size through the integrated studies.
2. To test if the effect sizes are homogeneous around the average effect size.
3. If the effect sizes are heterogeneous, to look for the influence of moderator variables through statistical techniques as the analysis of variance or regression analysis.

2. Procedure for averaging a set of independent effect sizes

- **We focus on the standardized mean difference, *d*, as the effect-size index in a meta-analysis. Let d_i to be the effect size of the i study in a meta-analysis:**

$$d_i = c(m)\frac{\bar{Y}_{Ei} - \bar{Y}_{Ci}}{S_i}$$

- **Assuming a random-effects model:**

d_i **estimates** $\theta_i \sim N(\mu, \tau^2)$

$$d_i = \theta_i + e_i$$

$$d_i \sim N(\mu, \tau^2 + v_i)$$

between-studies variance (heterogeneity variance)

within-study variance

2. Procedure for averaging a set of independent effect sizes

- **Assuming a random-effects model:**

d_i **estimates** [illegible]

[illegible]

[illegible]

between-studies variance (heterogeneity variance) — within-study variance

- **General procedure for averaging:**

$$\bar{d} = \frac{\sum w_i d_i}{\sum w_i}$$

[illegible]

heterogeneity variance estimator

3. Objective

➢ **To compare the statistical performance (bias and mean squared error) of nine procedures to average a set of independent effect sizes based on different heterogeneity variance estimators.**

$$\bar{d} = \frac{\sum w_i d_i}{\sum w_i} \qquad w_i = 1/(\hat{\sigma}_i^2 + \hat{\tau}^2)$$

- *Noniterative heterogeneity variance estimators:*

Hunter and Schmidt (1990): $$\hat{\tau}^2_{HS} = \frac{Q-k}{\sum w_i}$$

Hedges (1983): $$\hat{\tau}^2_{HE} = \frac{\sum_i (d_i - \bar{d}_U)^2}{k-1} - \frac{1}{k}\sum_i \hat{\sigma}_i^2$$

3. Objective

$$\bar{d} = \frac{\sum w_i d_i}{\sum w_i}$$

[illegible]

- *Noniterative heterogeneity variance estimators:*

Dersimonian and Laird (1986): $\hat{\tau}^2_{DL} = \frac{Q-(k-1)}{c}$

Malzahn et al. (2000): $\hat{\tau}^2_{MBH} = \frac{\sum_i (1-\phi_i)(d_i - \bar{d}_{m})^2}{k-1} - \frac{1}{k}\sum_i \left(\frac{N_i}{n_{Ei} n_{Ci}}\right) - \frac{1}{k}\sum_i \phi_i d_i^2$

[illegible]

3. Objective

$$\bar{d} = \frac{\sum w_i d_i}{\sum w_i}$$

[illegible]

- *Noniterative heterogeneity variance estimators:*

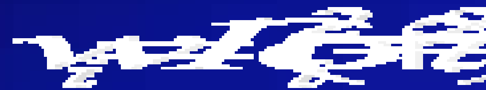
Hartum and Makambi (1986): $$\hat{\tau}^2_{HM} = \frac{Q^2}{[2(k-1)+Q]b}$$

Sidik and Jonkman (2005): $$\hat{\tau}^2_{SJ} = \frac{\sum_i \hat{v}_i^{-1}(d_i - \bar{d}_{\hat{v}})^2}{k-1}$$ $$\hat{v}_i = r_i + 1$$ $$r_i = \hat{\sigma}_i^2 / \hat{\tau}_0^2$$

$$\hat{\tau}_0^2 = \frac{\sum_i (d_i - \bar{d})^2}{k}$$ $$\bar{d}_{\hat{v}} = \frac{\sum_i \hat{v}_i^{-1} d_i}{\sum_i \hat{v}_i^{-1}}$$

3. Objective

$$\bar{d} = \frac{\sum w_i d_i}{\sum w_i}$$



- *Iterative heterogeneity variance estimators:*

Maximum likelihood:

$$\hat{\tau}^2_{ML} = \frac{\sum_i w_i^2\left[\left(d_i - \bar{d}_{ML}\right) - \hat{\sigma}_i^2\right]^2}{\sum_i w_i^2} \qquad \bar{d}_{ML} = \frac{\sum_i w_i d_i}{\sum_i w_i} \qquad w_i = 1/(\hat{\tau}^2 + \hat{\sigma}_i^2)$$

Restricted maximum likelihood:

$$\hat{\tau}^2_{REML} = \frac{\sum_i w_i^2\left[\left(d_i - \bar{d}_{ML}\right) - \hat{\sigma}_i^2\right]^2}{\sum_i w_i^2} + \frac{1}{\sum_i w_i}$$

3. Objective

$$\bar{d} = \frac{\sum w_i d_i}{\sum w_i}$$

[illegible]

- *Optimum average:*

$$\bar{d}_{Optimum} = \frac{\sum n_i d_i}{\sum n_i}$$

[illegible]

[illegible]

4. Background

- Viechtbauer (2005) compared the statistical performance of five out of our nine procedures to average a set of independent effect sizes, using as the heterogeneity variance estimator: $\hat{\tau}^2_{HS}$ (Hunter & Schmidt, 1990), $\hat{\tau}^2_{HE}$ (Hedges 1983), $\hat{\tau}^2_{DL}$ (DerSimonian & Laird, 1986), $\hat{\tau}^2_{ML}$ (maximum likelihood estimator, Brockwell & Gordon, 2001), and $\hat{\tau}^2_{REML}$ (restricted maximum likelihood estimator, Brockwell & Gordon, 2001).

- **First conclusion:** Estimates of μ_δ were slightly negatively biased, the bias being the same regardless of which heterogeneity variance estimator was used to compute the average (p. 285).

- **Second conclusion:** Calculating the w_i values using unbiased estimates of τ^2 and σ^2_{di} results in an estimate of the sampling variance of $\bar{d}$ that is negatively biased (p. 263).

$$\bar{d} = \frac{\sum w_i d_i}{\sum w_i}$$

5. Method

- The effect-size index was the standardized mean difference, d_i:

$$d_i = c(m_i)\frac{\bar{Y}_{Ei} - \bar{Y}_{Ci}}{S_i}$$

- With estimated within-study sampling variance, $\hat{\sigma}_i^2$:

$$\hat{\sigma}_i^2 = \frac{n_{Ei} + n_{Ci}}{n_{Ei} n_{Ci}} + \frac{d_i^2}{2(n_{Ei} + n_{Ci})}$$

- A random-effects model was assumed, where $\delta_i \sim N(\mu_\delta, \tau^2)$:

$$\delta_i = \frac{\mu_{Ei} - \mu_{Ci}}{\sigma_i}$$

- To define each parametric effect δ_i, two normal distributions, $N(\mu_{Ei}, \sigma_i^2)$ and $N(\mu_{Ci}, \sigma_i^2)$, were assumed

5. Method

- Pairs of independent samples with $n_E = n_C$ units were generated from these populations. Each pair of samples simulated the data of a primary study.
- The value of the parametric mean effect size was always set at $\mu_\delta = 0.5$
- Manipulated factors:

  Variability of δ_i: $\tau^2 = 0, .04, .08, .16$, and $.32$
  Number of studies of the meta-analysis: $k = 5, 10, 20, 40$, and 100
  Average sample size of the studies in a meta-analysis: $\bar{N} = 30, 50, 80, 100$
- 10,000 replies (meta-analyses) for each of the 100 (5x4x5) conditions were simulated

5. Method

- ♦ Computations in each simulated meta-analysis:

  - The standardized mean difference (*d* of Hedges) in each study

  - The nine averages, $\bar{d}$, one for each τ^2 estimator:

$$\bar{d}_{HS} \quad \bar{d}_{HE} \quad \bar{d}_{DL} \quad \bar{d}_{MBH} \quad \bar{d}_{HM} \quad \bar{d}_{SJ} \quad \bar{d}_{ML} \quad \bar{d}_{REML} \quad \bar{d}_{OPTIM}$$

$$\bar{d} = \frac{\sum w_i d_i}{\sum w_i}$$

[illegible]

5. Method

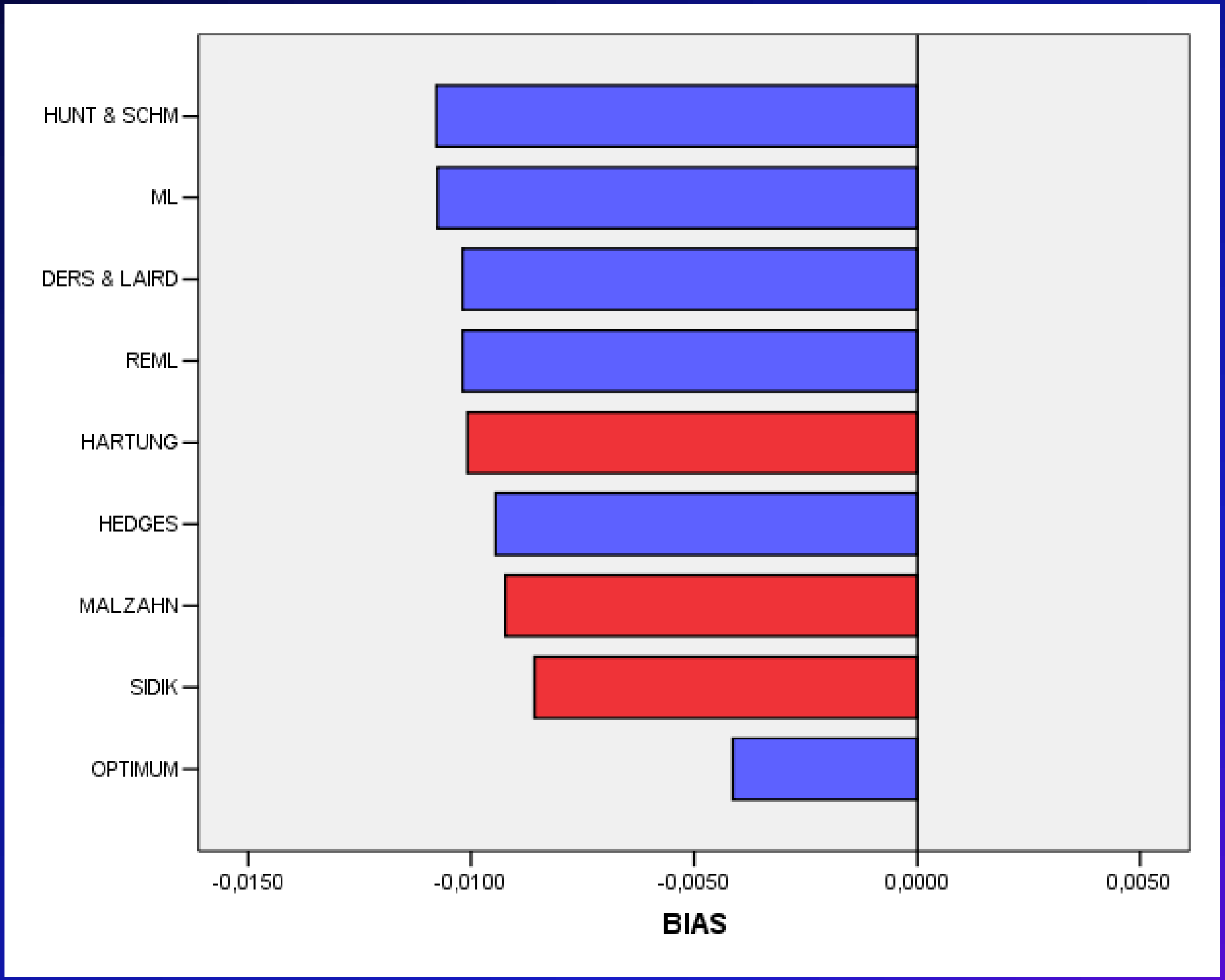
♦ Estimations through the 10,000 replies of a same meta-analysis

- Bias and Mean Squared Error of the nine averages:

$$\text{Bias}\left(\bar{d}\right) = \frac{\sum d_i}{10{,}000} - \mu_\delta$$

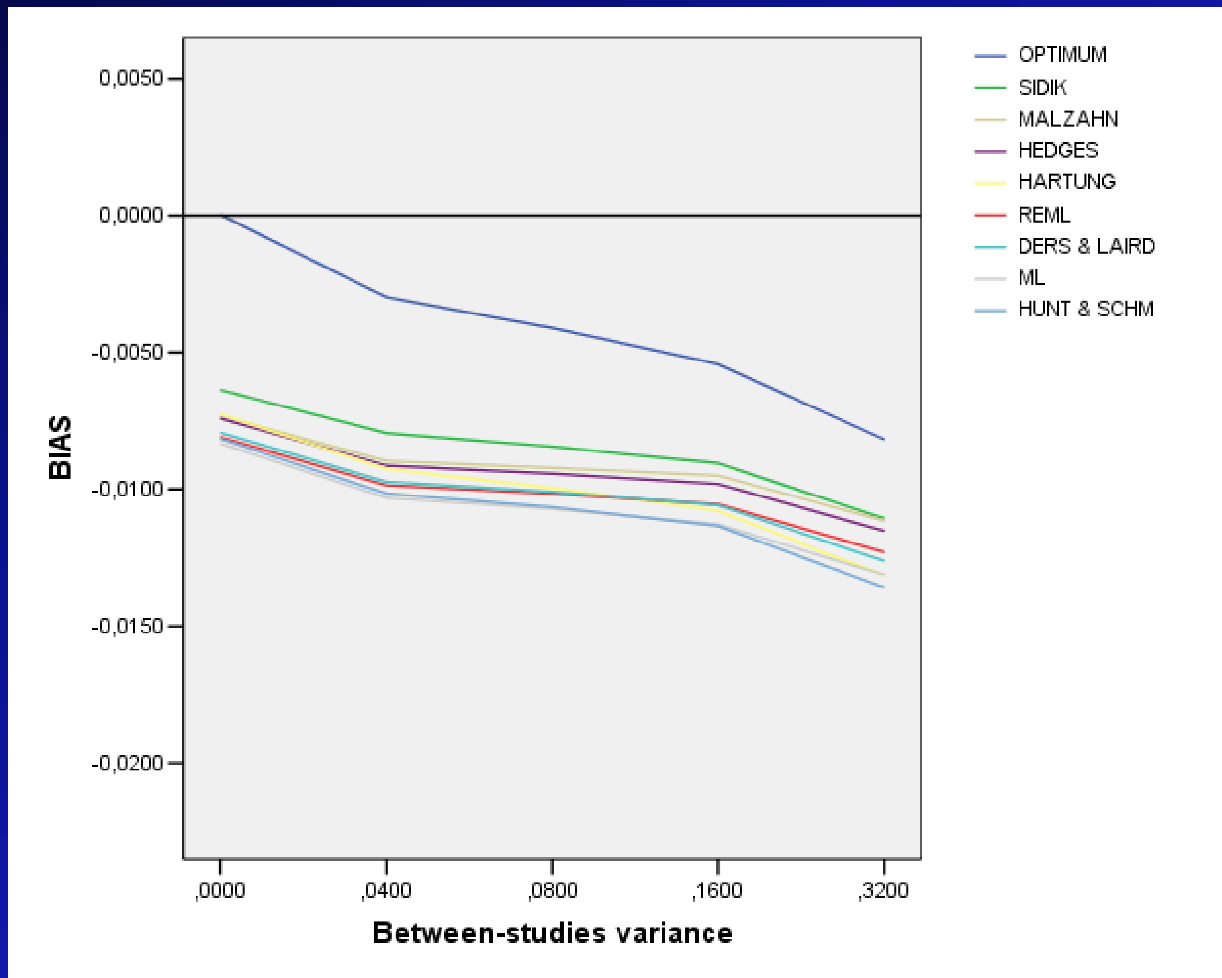
$$\text{MSE}\left(\bar{d}\right) = \frac{\sum \left(d_i - \mu_\delta\right)^2}{10{,}000}$$

6. Results: BIAS of the nine averages

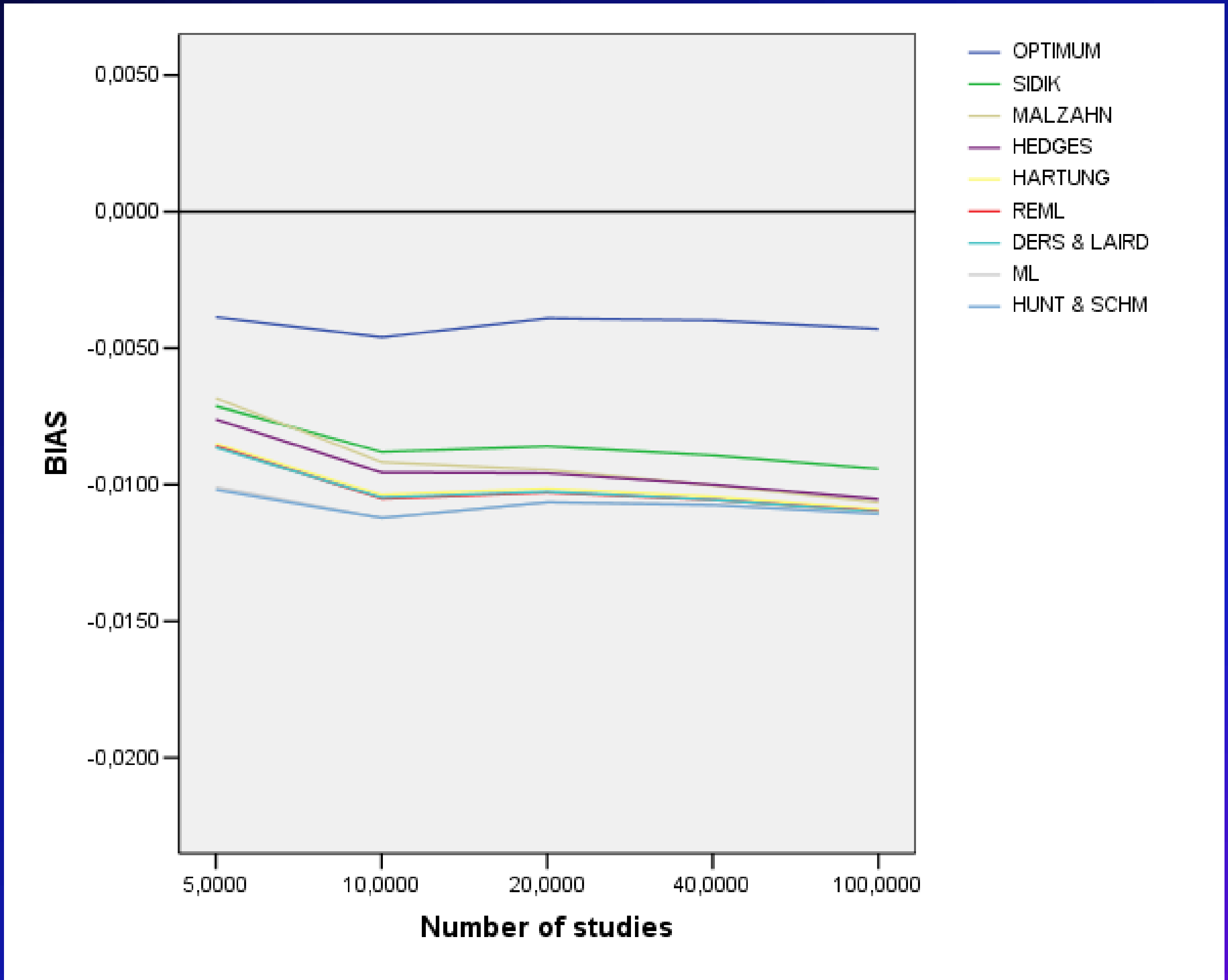


Sidik, Malzahn and Hedges procedures are the closest to the optimum one

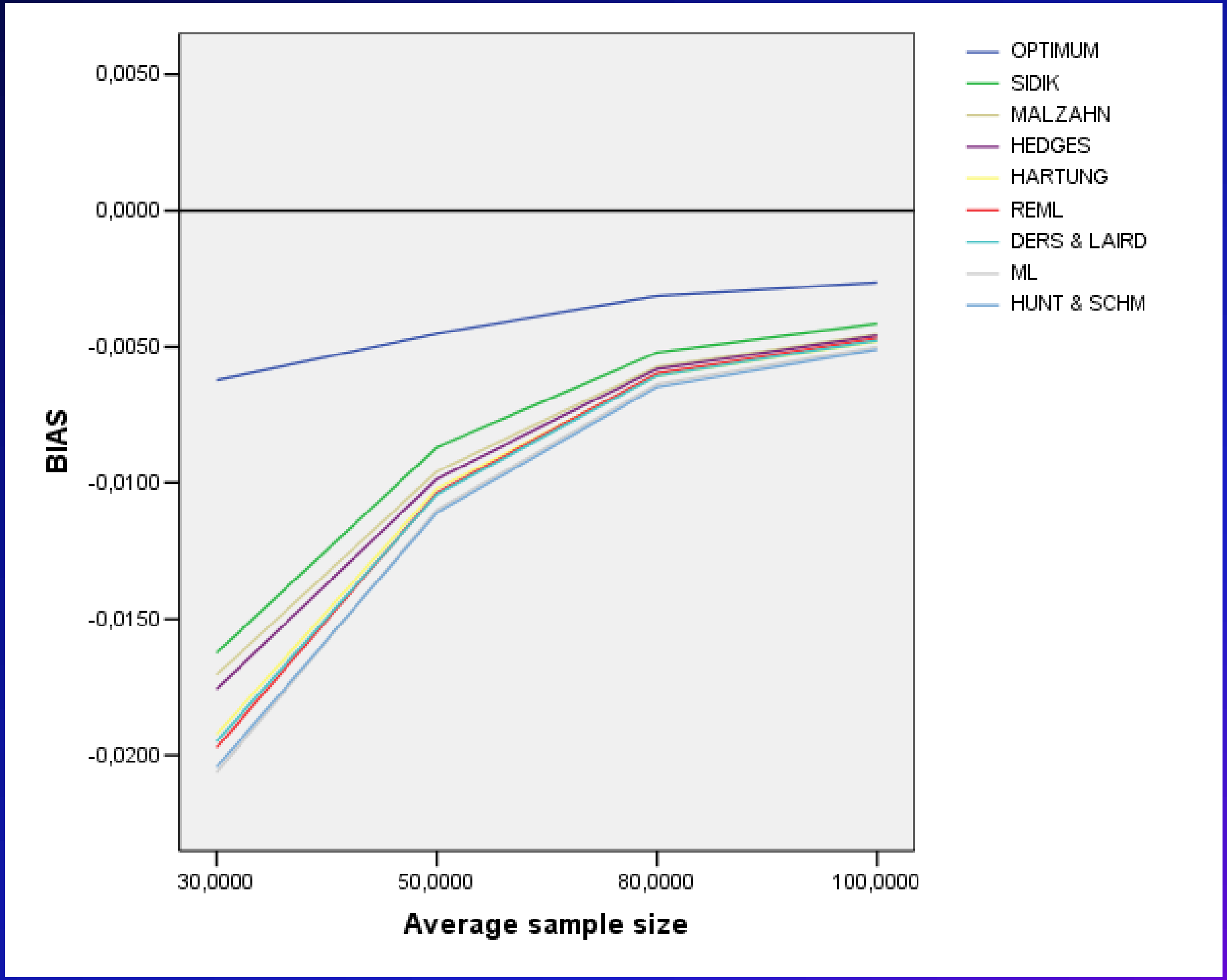
6. Results: BIAS as a function of the between-studies variance



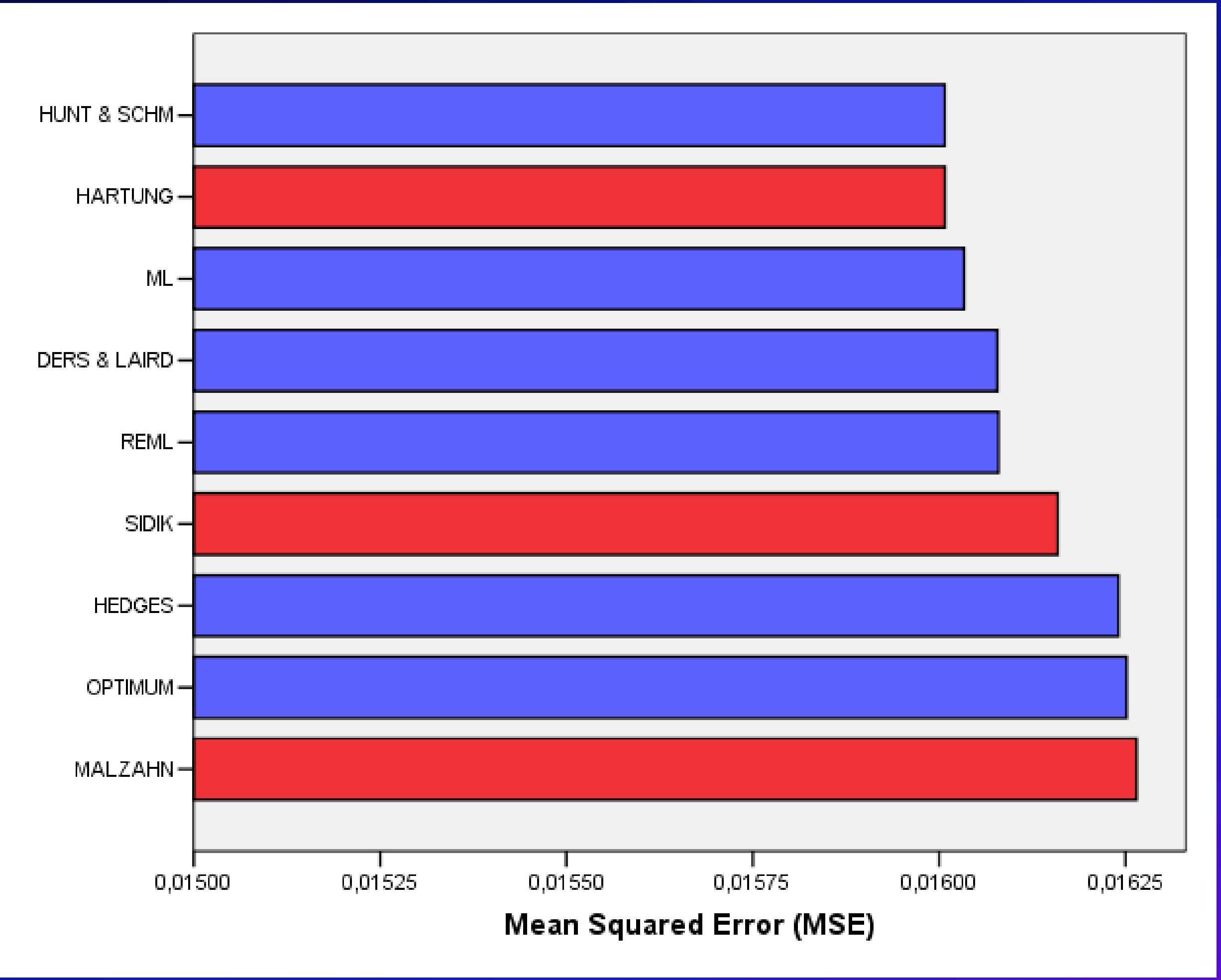
6. Results: BIAS as a function of the number of studies



6. Results: BIAS as a function of the sample size

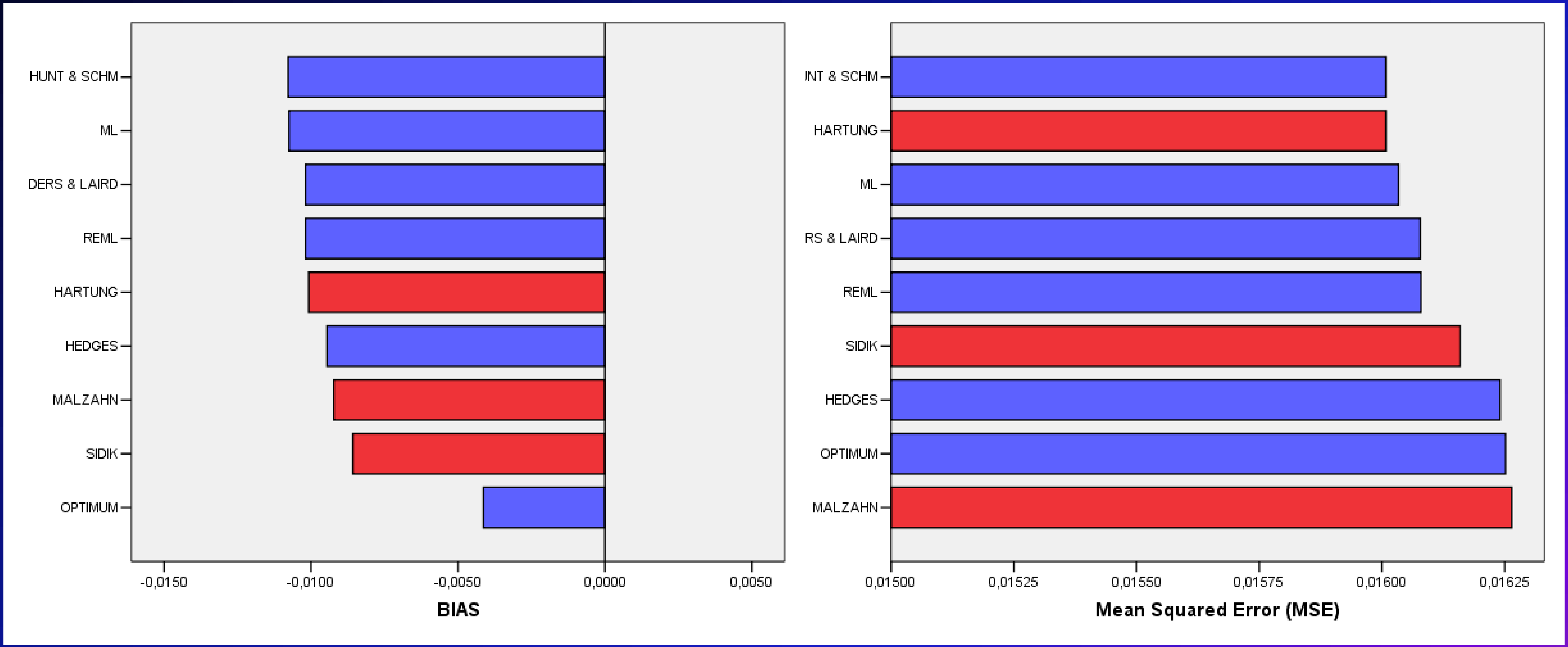


6. Results: Mean Squared Error of the nine averages



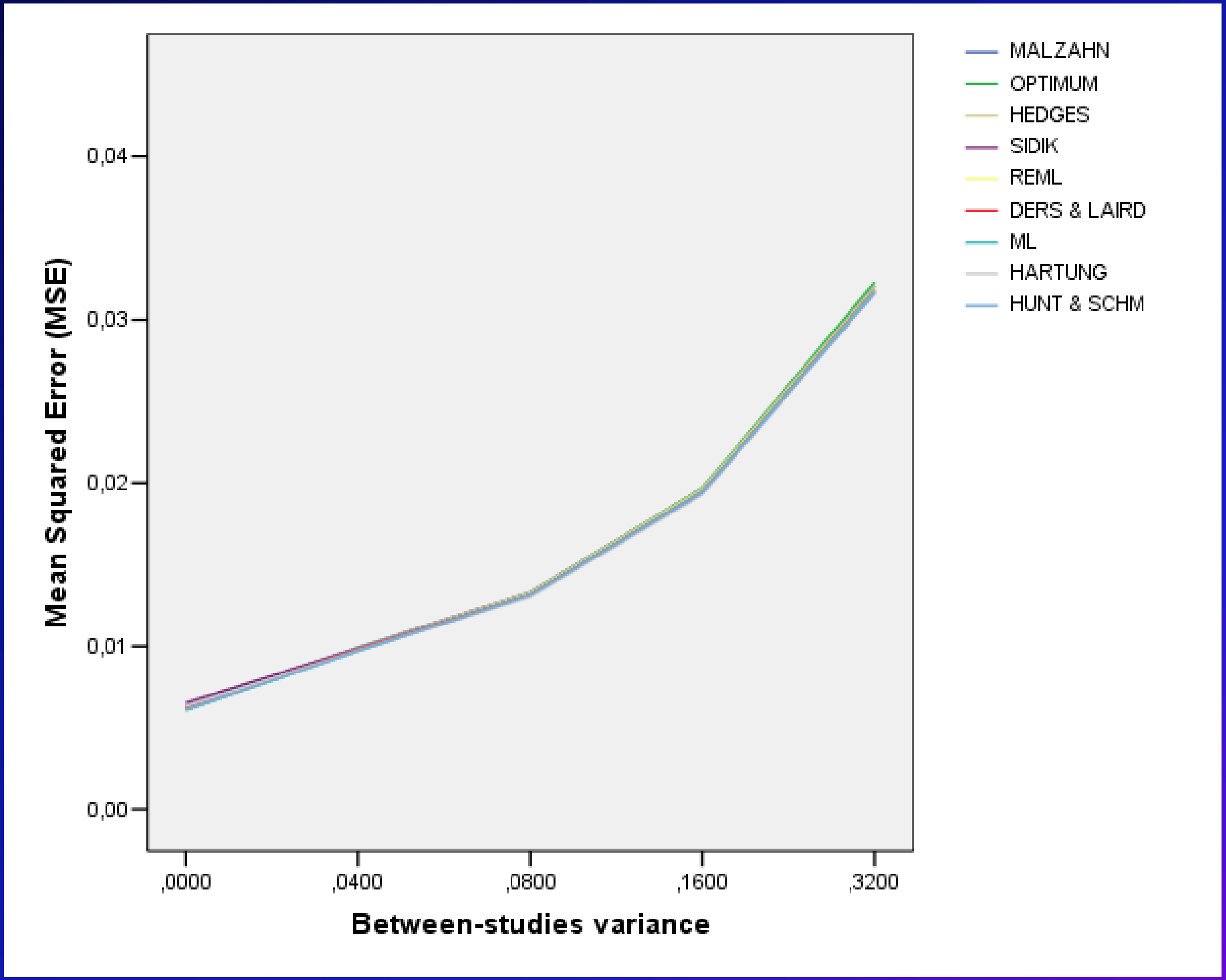
Sidik, Malzahn and Hedges procedures are the closest to the optimum one

6. Results: BIAS and MSE of the nine averages

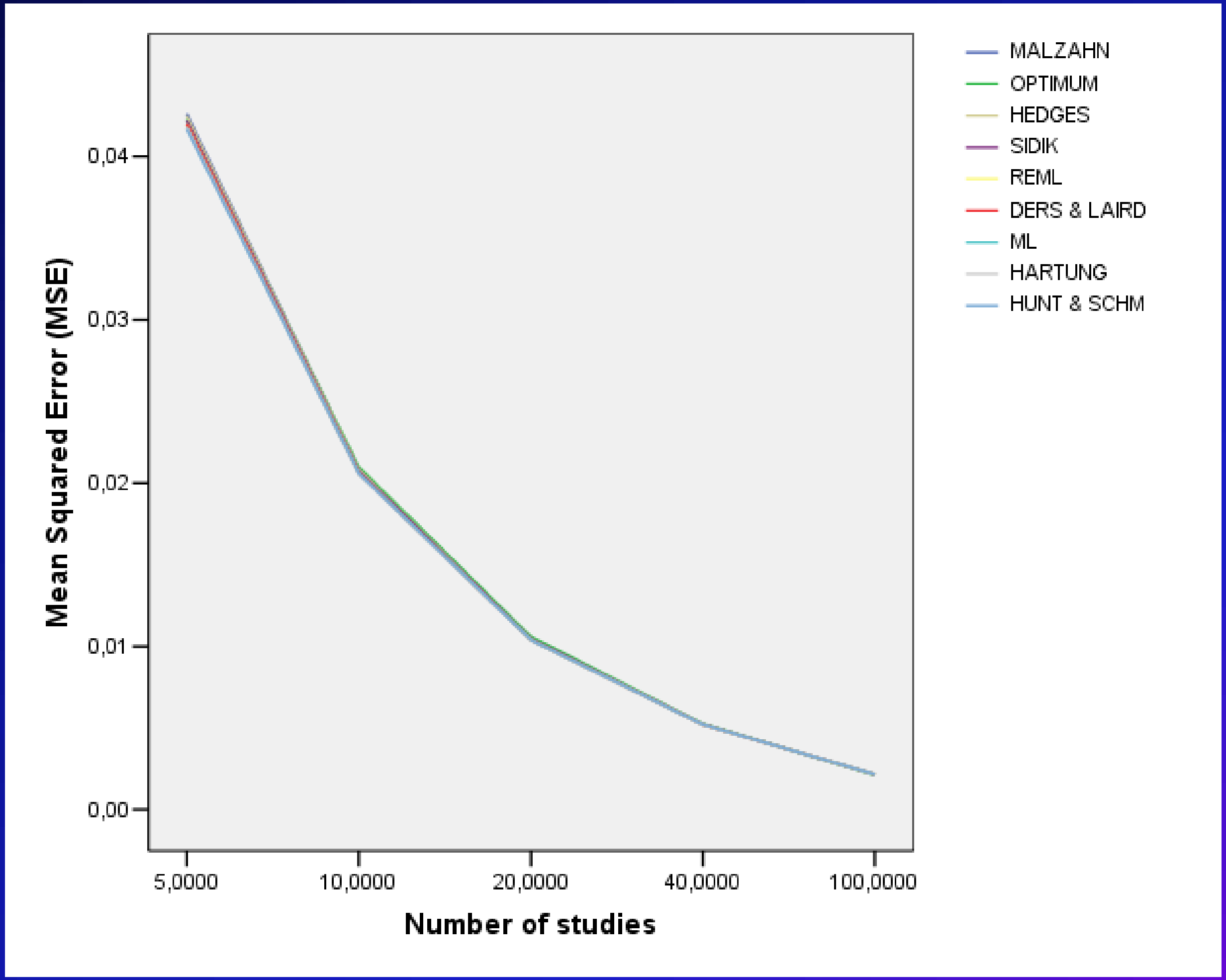


Sidik, Malzahn and Hedges procedures are the closest to the optimum one

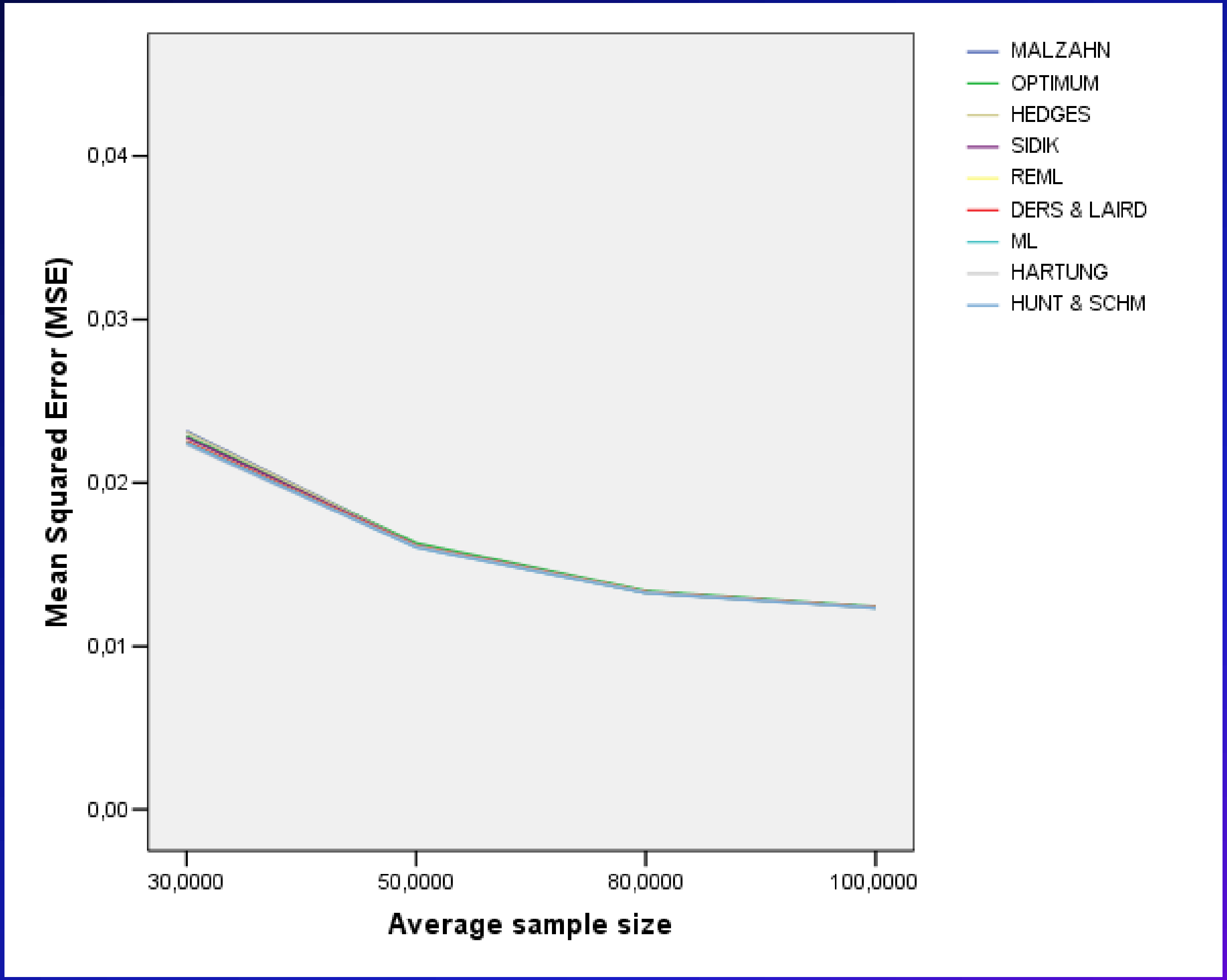
6. Results: MSE as a function of the between-studies variance



6. Results: MSE as a function of the number of studies



6. Results: MSE as a function of the sample size



7. Conclusions

1. The selection among different heterogeneity variance estimators slightly affects to the bias and mean squared error of the overall effect size in a meta-analysis.
2. All of the procedures to estimate an average effect size, with the different heterogeneity variance estimators, coincided in showing a slight negative bias. The largest bias values (around the 4%) were found in meta-analyses with small sample sizes and a large between-studies variance.
3. With the exception of the Malzhan's procedure, the other ones slightly underestimated the mean squared error showed in the optimum average.
4. Taking into consideration both the bias and the mean squared error, the procedures of Sidik & Jonkman (2005), Malzahn et al. (2000) and Hedges (1983) showed the best performance.
5. The conclusions of our results are limited to the simulated conditions and the use of the standardized mean difference as the effect size index in a meta-analysis.

MAY 14-16, 2007

LONDON, ENGLAND

The Seventh Annual International Campbell Collaboration Colloquium

Quality, Credibility, and Utility: The Relevance of Systematic Reviews

Symposium: Advances in Meta-analysis

Coordinator: Julio Sánchez Meca

Estimating an Overall Effect Size: Does the Between-Studies Variance Estimator have an Effect?

Fulgencio Marín Martínez and Julio Sánchez Meca

Universidad de Murcia

http://www.um.es/facpsi/metaanalysis

This research was supported by a Grant of the Ministerio de Educación y Ciencia and by FEDER Founds (Project nº: SEJ2004-0727B).