Small study effects in meta-analysis

Gerta Rücker, Guido Schwarzer[1]

[1] Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center – University of Freiburg, Germany
e-mail: ruecker@imbi.uni-freiburg.de

DFG research project RU1747/1-2

Mastercourse Biometry/Biostatistics, 20. June 2020

Outline

1. **Definition of small-study effects**
2. **Diagnosis of small-study effects**
3. **Adjusting for small-study effects**

Bias: Overview

- Types of bias
  - Publication bias and related biases
  - Small-study effects
- Diagnosis of small-study effects
  - Funnel plot
  - Funnel plot tests
- Adjustment for small-study effects
  - Trim and fill method
  - Copas selection model
  - Adjustment by regression

Bias in meta-analyses: Small-study effects

- Publication bias [Easterbrook et al., 1991, Rothstein et al., 2005]: Small studies tend to be published only if they show a large effect
- Related types of bias: Studies having 'significant' results tend to be
  - published in high-ranking English language journals *(Language bias)* [Egger et al., 1997b]
  - published faster than studies without a 'significant' result *(Time lag bias)* [Higgins and Green, 2009]
  - published more than once *(Multiple publication bias)* [Gøtzsche, 1989]
  - cited more often than studies without a 'significant' result, and therefore are more easily detectable in literature searches *(Citation bias)* [Nieminen et al., 2007]

Small-study effects

Smaller trials show different, often larger, treatment effects than large ones [Sterling et al., 1995, Sterne et al., 2000, Rothstein et al., 2005]

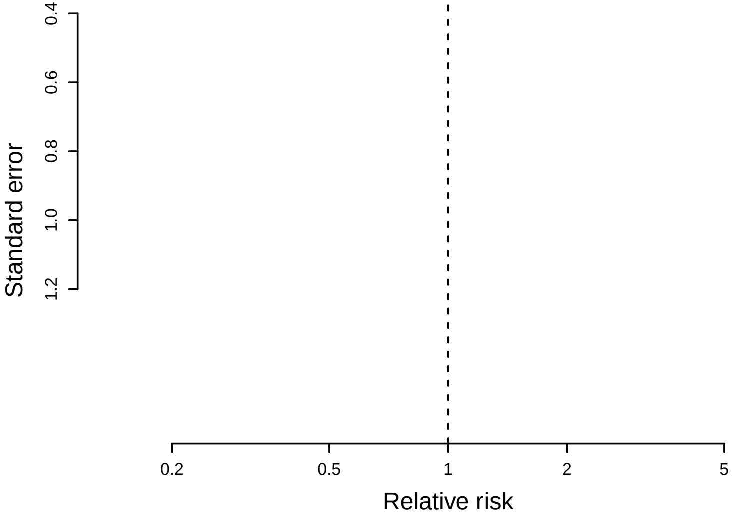
- Potential causes of small-study effects
  - Publication bias: Small studies tend to be published preferably if they show a large effect [Easterbrook et al., 1991]
  - Selective outcome reporting bias: Studies present selected outcomes [Chan et al., 2004a, Chan et al., 2004b, Williamson and Gamble, 2005]
  - Selective analysis reporting bias: Studies choose a method of analysis that leads to larger effects [Ioannidis et al., 2014]
  - Clinical heterogeneity between patients in large and small trials
  - For binary data: Statistical correlation between treatment effect estimate and its variance [Schwarzer et al., 2002]
  - Coincidence
- Graphical representation of small-study effects
  - Funnel plot [Light and Pillemer, 1984, Sterne and Egger, 2001]

Funnel plot

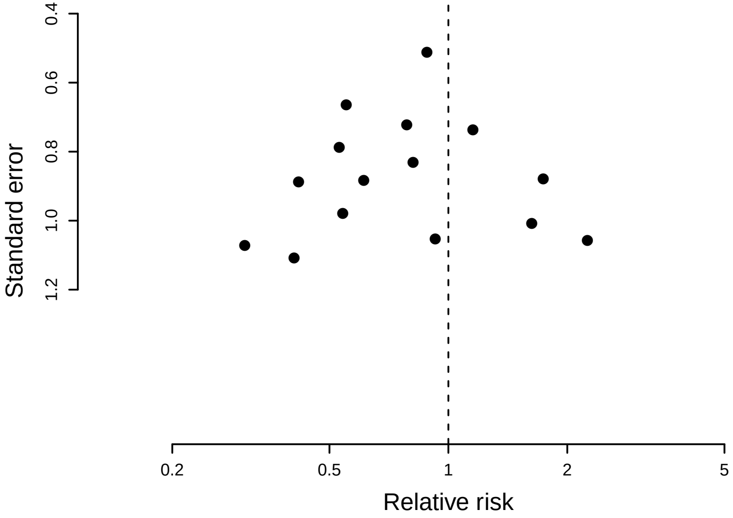
Horizontal axis: (log) treatment effect

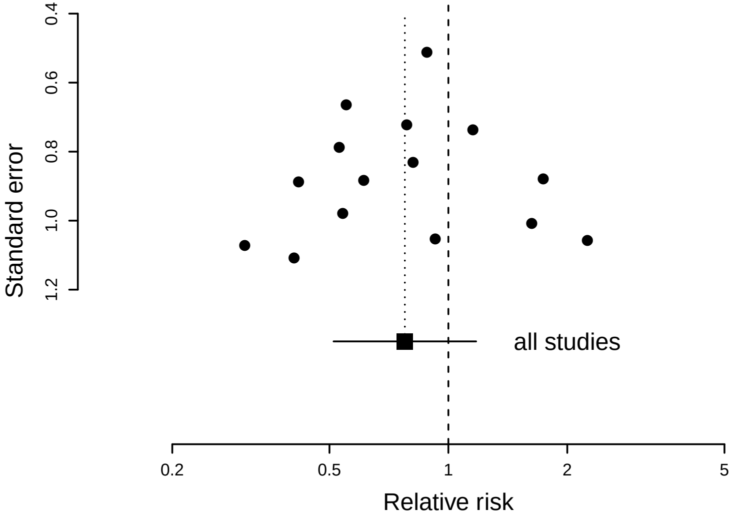
Vertical axis: a measure of precision; various versions in the literature:

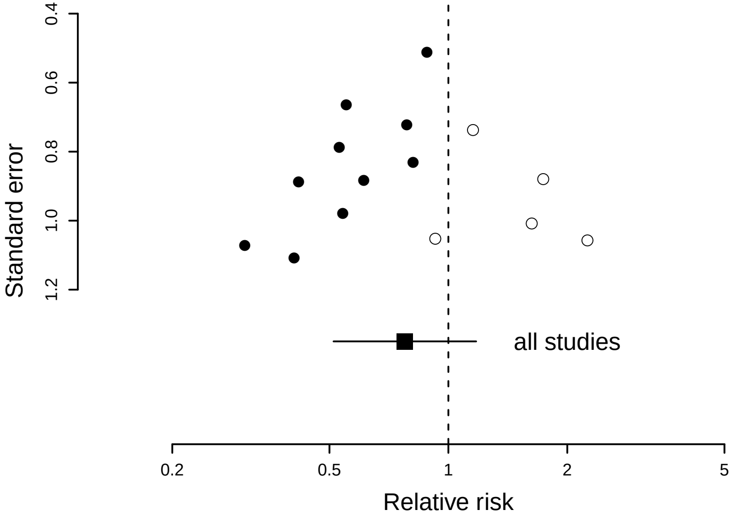
- No longer recommended: Sample size, Inverse variance
- Preferred: Standard error on a reversed axis [Sterne and Egger, 2001]
    - confidence intervals increasing linearly
    - sufficient space for imprecise (small) studies (particularly interesting for diagnosis of small-study effects)

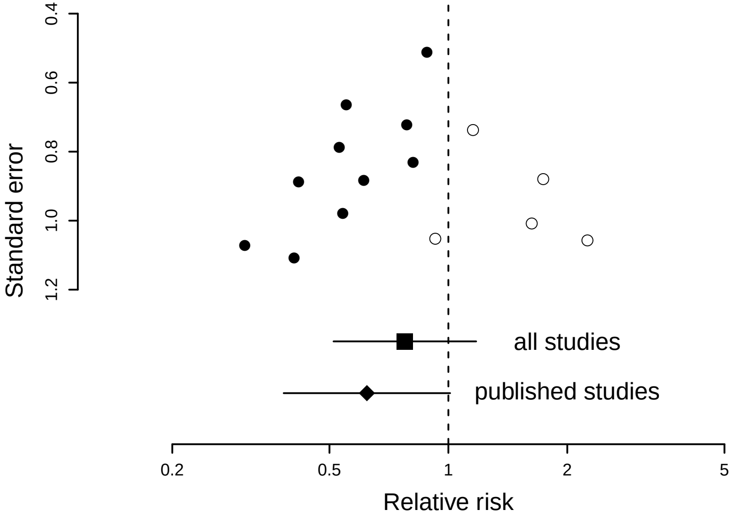


Standard error
0.4
0.6
0.8
1.0
1.2
0.2
0.5
1
2
5
Relative risk


Standard error
0.4
0.6
0.8
1.0
1.2
0.2
0.5
1
2
5
Relative risk

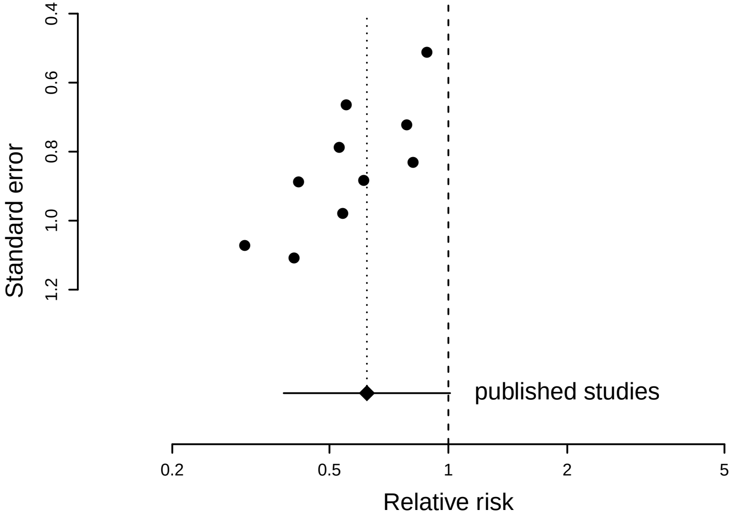

Standard error
0.4
0.6
0.8
1.0
1.2
all studies
0.2
0.5
1
2
5
Relative risk


Standard error
0.4
0.6
0.8
1.0
1.2
all studies
0.2
0.5
1
2
5
Relative risk



Standard error
0.4
0.6
0.8
1.0
1.2
all studies
published studies
0.2
0.5
1
2
5
Relative risk

Asymmetry


Standard error
0.4
0.6
0.8
1.0
1.2
published studies
0.2
0.5
1
2
5
Relative risk

Funnel plot

Example: NSAIDS data

Meta-analysis of 37 placebo-controlled randomized trials on the effectiveness and safety of topical non-steroidal anti-inflammatory drugs (NSAIDS) in acute pain [Moore et al., 1998]

Part of R package metasens

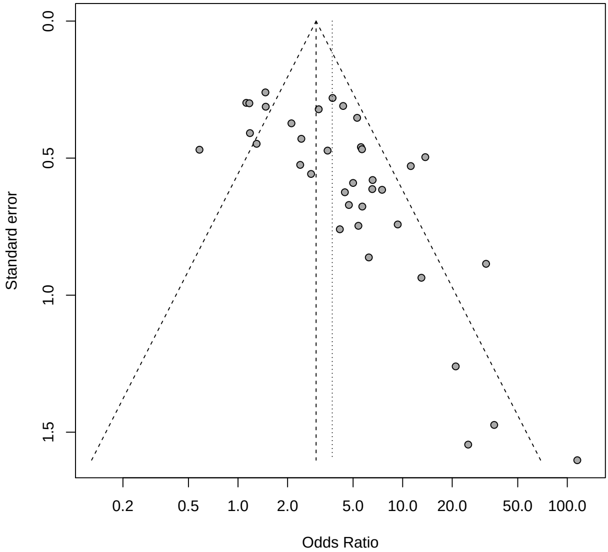
How to obtain a funnel plot in R

```
# The data are part of library metasens, therefore this must be loaded
library(metasens)

# Load the data and look at the variable names
data(nsaids)
names(nsaids)

## [1] "study" "Ee"    "Ne"    "Ec"    "Nc"
```

```
# Perform meta-analysis
ms1 <- metabin(Ee, Ne, Ec, Nc, data = nsaids, sm = "OR")
```

```
# Create funnel plot
funnel(ms1)
```

Funnel plot of NSAIDS data



Compare results of common effect and random effects model

```
summary(ms1)

## Number of studies combined: k=37
##
##                            OR           95%-CI       z  p-value
## Fixed effect model   2.9809 [2.5854; 3.4368] 15.0409 < 0.0001
## Random effects model 3.7345 [2.8039; 4.9740]  9.0105 < 0.0001
##
## Quantifying heterogeneity:
## tau^2 = 0.4670; H = 1.78 [1.5; 2.1]; I^2 = 68.3% [55.5%; 77.4%]
##
## Test of heterogeneity:
##       Q d.f.  p-value
##  113.52   36 < 0.0001
##
## Details on meta-analytical method:
## - Mantel-Haenszel method
## - DerSimonian-Laird estimator for tau^2
## - Continuity correction of 0.5 in studies with zero cell frequencies
```

Funnel plot-based tests for small-study effects

- Idea: Test for asymmetry in the funnel plot as an indication for bias
- Method: Test for association between treatment effect and standard error
- Assumption: No association between treatment effect and standard error (or precision) if there is no small-study effect
- Limitation: Strictly valid only for normally distributed outcomes
- Criticised by some authors [Terrin et al., 2005, Lau et al., 2006]
  - Other types of tests, not available in R package meta: [Vevea and Hedges, 1995, Hedges and Vevea, 1996, Ioannidis and Trikalinos, 2007]

Funnel plot tests for asymmetry: Overview

Rank correlation tests (not considered here)

- [Begg and Mazumdar, 1994]
  - Modification for binary data: [Schwarzer et al., 2007]

Regression tests

- [Egger et al., 1997a]
- Modifications for binary data
  - [Harbord et al., 2006]
  - [Macaskill et al., 2001]
  - [Peters et al., 2006]
  - Arcsine test [Rücker et al., 2008]

Regression tests: Basic idea

Choose an effect measure, say, the mean difference

Null-hypothesis (*‘No small study effects’*): Treatment effect does not depend on precision

1. Regress the treatment effect on the standard error, using inverse variance weights
2. Test null-hypothesis of zero slope

Often called Egger’s test [Egger et al., 1997a]

Note: Strictly valid only for continuous data (data normally distributed)!

Nevertheless often applied to binary data, preferably in a modified version

How to obtain a regression test in R

```r
# Perform Egger's test using R function metabias
metabias(ms1, method = "linreg")
```

How to obtain a regression test in R

```r
# Perform Egger's test using R function metabias
metabias(ms1, method = "linreg")

##
##  Linear regression test of funnel plot asymmetry
##
## data:  ms1
## t = 4.7147, df = 35, p-value = 3.786e-05
## alternative hypothesis: asymmetry in funnel plot
## sample estimates:
##       bias    se.bias      slope
##  2.7652744  0.5865197 -0.1122134
```

Modifications of Egger's test for binary data

- Harbord's score test [Harbord et al., 2006]
  - Uses a score-based estimate for the odds ratio
  - Advantage: Variance estimate depends only on marginal totals
  - Use R: `metabias(ms1, method = "score")`
- Peters' test [Peters et al., 2006]
  - Uses the usual odds ratio estimate and $1/n$ as regressor
  - Advantage: Study weights depending only on marginal totals
  - Use R: `metabias(ms1, method = "peters")`
- Arcsine test [Rücker et al., 2008]
  - Uses the arcsine difference instead of the odds ratio
  - Advantage: Variance depends only on group sample sizes
  - Use R:

```
ms1.asd <- update(ms1, sm = "ASD")
metabias(ms1.asd, method = "linreg")
```

Recommendations on testing for funnel plot asymmetry

[Sterne et al., 2011], BMJ:

- Funnel plot tests only when there are at least 10 studies (rule of thumb; argument `k.min` in function `metabias`)
- Recommendation for continuous outcomes:
  Linear regression test [Egger et al., 1997a]
- For binary outcomes:
  Use one of the modifications of Egger's test
  [Harbord et al., 2006, Peters et al., 2006, Rücker et al., 2008]
- Bias cannot be excluded if test for funnel plot asymmetry is non-significant
- Test performance deteriorates if between-study heterogeneity increases

Adjusting for small-study effects

Three approaches

- **Trim and fill method**
  [Duval and Tweedie, 2000a, Duval and Tweedie, 2000b]
- **Copas selection model for publication bias**
  [Copas, 1999, Copas and Shi, 2000, Copas and Shi, 2001]
- **Adjustment by regression**
  [Copas and Malley, 2008, Stanley, 2008, Moreno et al., 2009a, Moreno et al., 2009b, Rücker et al., 2011b, Rücker et al., 2011a]

Trim and fill method

1. Estimate the number of studies in the outlying part of the funnel plot using rank-based methods;
2. remove (trim) these studies and do meta-analysis on the remaining studies;
3. consider the estimate from the 'trimmed' meta-analysis as the true center of the funnel;
4. for each 'trimmed' study, create (fill) an additional study as the mirror image about the center of funnel plot;
5. do meta-analysis on original and filled studies.

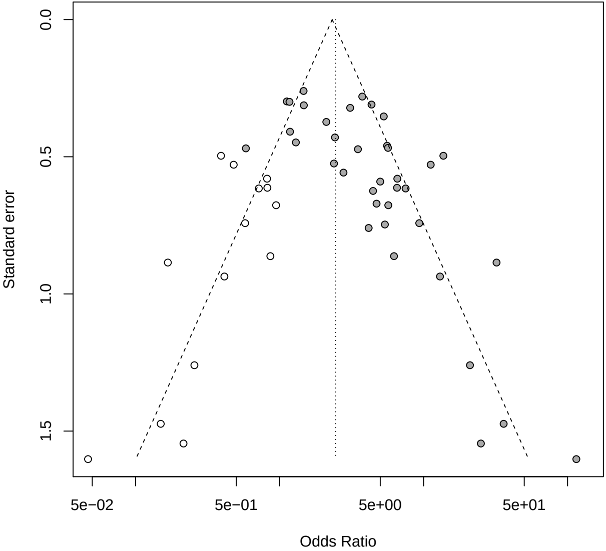
How to perform a Trim and fill analysis in R

```
# Conduct meta-analysis
ms1 <- metabin(Ee, Ne, Ec, Nc, data = nsaids, sm = "OR")
```

```
# Perform Trim and fill analysis
tf1 <- trimfill(ms1)
```

```
# Create funnel plot including filled-in studies
funnel(tf1)
```

Trim and fill plot of NSAIDS data



How to perform a Trim and fill analysis in R

```
# Print results of Trim and fill analysis
print(tf1, digits = 2)

##                OR           95%-CI %W(random)
## 1            6.57 [2.11;   20.48]       2.13
*** Output truncated ***
## 37           5.69 [1.51;   21.42]       1.91
## Filled: 37   0.95 [0.25;    3.56]       1.91
## Filled: 27   0.86 [0.16;    4.68]       1.53
## Filled: 16   0.82 [0.25;    2.73]       2.05
*** Output truncated ***
## Filled: 32   0.05 [0.00;    1.08]       0.68
##
## Number of studies combined: k=51 (with 14 added studies)
##
##                         OR       95%-CI z  p-value
## Random effects model 2.45 [1.83; 3.28] 6 < 0.0001
##
## Quantifying heterogeneity:
## tau^2 = 0.7113; H = 1.93 [1.68; 2.22]; I^2 = 73.2% [64.7%; 79.7%]
*** Output truncated ***
```

Copas selection model

Combine two models:

1. Usual random effects model for treatment effect
2. A model for the selection process with a parameter controlling how chance of publication depends on precision $1/s_k$ (where s_k is the within-study standard error)

Selection/publication bias is modelled by a parameter representing the correlation between effect size and selection probability

- Implemented in function `copas` of R package metasens (earlier: copas) [Carpenter et al., 2009a]
- Sensitivity analysis necessary
- Not treated here in detail

Adjustment by regression

- Random effects model

$$\hat{\theta}_k = \theta_k + \sigma_k\, \eta_k, \qquad \eta_k \sim N(0,1)$$
$$\theta_k = \theta + \tau\, \delta_k, \qquad \delta_k \sim N(0,1)$$

  - $\hat{\theta}_k$ observed treatment effect in study k $(k = 1, \dots, K)$
  - θ_k true treatment effect in study k
  - θ overall treatment effect
  - σ_k^2 within-study sampling variance, τ^2 between-study variance

- Equivalent:

$$\hat{\theta}_k = \theta + \sqrt{\sigma_k^2 + \tau^2}\, \epsilon_k, \qquad \epsilon_k \sim N(0,1)$$

Adjustment by regression

- Random effects model:

$$\hat{\theta}_k = \theta + \sqrt{\sigma_k^2 + \tau^2}\, \epsilon_k, \qquad \epsilon_k \sim N(0,1)$$

- Extended random effects model taking account of possible small study effects by allowing the effect to depend on the standard error:

$$\hat{\theta}_k = \theta + \sqrt{\sigma_k^2 + \tau^2}\; (\alpha + \epsilon_k), \qquad \epsilon_k \sim N(0,1)$$

  Additional parameter α represents bias introduced by small-study effects ('publication bias')

Adjustment by regression

- Extended random effects model

$$\hat{\theta}_k = \theta + \sqrt{\sigma_k^2 + \tau^2}\ (\alpha + \epsilon_k), \qquad \epsilon_k \sim N(0,1)$$

- Additional parameter α represents bias introduced by small-study effects ('publication bias')
  - For a very small study k, we have $\sigma_k^2 \to \infty$ and therefore

$$\mathrm{E}\left(\frac{\hat{\theta}_k - \theta}{\sigma_k}\right) \to \alpha \quad \text{Small study bias}$$

  - For a very large study k, we have $\sigma_k^2 \to 0$ and therefore

$$\mathrm{E}\left(\hat{\theta}_k\right) \to \theta + \tau\,\alpha \quad \text{Adjusted effect of large study}$$

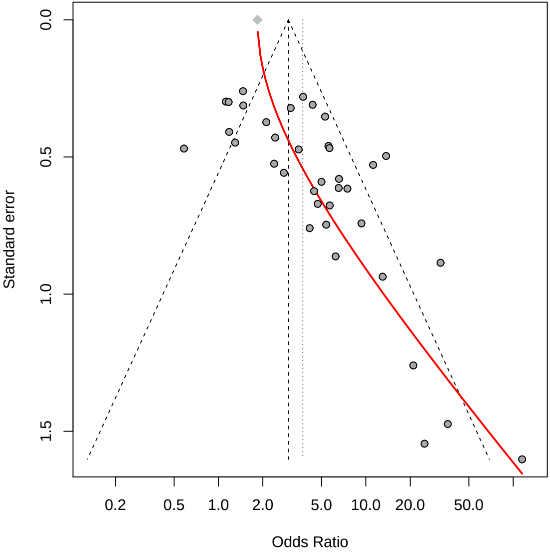
- Implemented in function `limitmeta` of R package metasens [Carpenter et al., 2009a]

How to perform regression adjustment in R

```r
# Perform limit meta-analysis
l1 <- limitmeta(ms1)
```

```r
# Create funnel plot with adjusted regression line
funnel(l1, col.line = "red", lwd.line = 2)
```

Funnel plot with adjusted regression line for NSAIDS data



How to perform regression adjustment in R

```
# Print results of regression adjustment (limit meta-analysis)
print(summary(l1), digits = 2)

## Result of limit meta-analysis:
##
##  Random effects model   OR       95%-CI    z     pval
##     Adjusted estimate 1.84 [1.26; 2.68] 3.17   0.0015
##   Unadjusted estimate 3.73 [2.80; 4.97] 9.01 < 0.0001
##
## Quantifying heterogeneity:
## tau^2 = 0.4670; I^2 = 68.3% [55.5%; 77.4%]; G^2 = 91.5%
##
## Test of heterogeneity:
##       Q d.f.  p.value
##  113.52   36 < 0.0001
##
## Test of small-study effects:
##    Q-Q' d.f.  p.value
##   44.20    1 < 0.0001
##
## Test of residual heterogeneity beyond small-study effects:
##      Q' d.f.  p.value
##   69.32   35   0.0005
##
```

Compare estimates for NSAIDS example

Model	Odds ratio [95% CI]
Common effect model	2.89 [2.49; 3.35]
Random effects model	3.73 [2.80; 4.97]
Trim and fill (random effects estimate)	2.45 [1.83; 3.28]
Copas selection model	1.82 [1.46; 2.26]
Regression adjustment	1.84 [1.26; 2.68]

Adjusting for small-study effects: Summary

Three approaches

- Trim and fill method
  - Easily conducted using R function `trimfill` in meta
  - Not model-based, somewhat ad hoc
- Copas selection model for publication bias
  - Model-based, needs sensitivity analysis
  - Function `copas`, implemented in R package metasens
  - Sometimes associated with estimation problems [Carpenter et al., 2009b]
- Adjustment by regression
  - Model-based, extension of the regression test
  - Function `limitmeta`, implemented in R package metasens

References I

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

Carpenter, J., Rücker, G., and Schwarzer, G. (2009a).
copas: An R package for fitting the Copas selection model.
The R Journal, 1(2):31–36.

Carpenter, J. R., Schwarzer, G., Rücker, G., and Künstler, R. (2009b).
Empirical evaluation showed that the Copas selection model provided a useful summary in 80% of meta-analyses.
Journal of Clinical Epidemiology, 62:624–631.

Chan, A.-W., Hróbartsson, A., Haahr, M. T., Gøtzsche, P. C., and Altman, D. G. (2004a).
Empirical evidence for selective reporting of outcomes in randomized trials. comparison of protocols to published articles.
Journal of the American Medical Association, 291:2457–2465.

Chan, A.-W., Krleza-Jeric, K., Schmid, I., and Altman, D. G. (2004b).
Outcome reporting bias in randomized trials funded by the Canadian institutes of health research.
Canadian Medical Association Journal, 171:735–740.

Copas, J. (1999).
What works?: Selectivity models and meta-analysis.
Journal of the Royal Statistical Society, Series A, 162:95–109.

Copas, J. and Shi, J. Q. (2000).
Meta-analysis, funnel plots and sensitivity analysis.
Biostatistics, 1:247–262.

Copas, J. B. and Malley, P. F. (2008).
A robust p-value for treatment effect in meta-analysis with publication bias.
Statistics in Medicine, 27(21):4267–4278.

References II

Copas, J. B. and Shi, J. Q. (2001).
A sensitivity analysis for publication bias in systematic reviews.
Statistical Methods in Medical Research, 10:251–265.

Duval, S. and Tweedie, R. (2000a).
A nonparametric “Trim and Fill” method of accounting for publication bias in meta-analysis.
Journal of the American Statistical Association, 95:89–98.

Duval, S. and Tweedie, R. (2000b).
Trim and Fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis.
Biometrics, 56:455–463.

Easterbrook, P. J., Berlin, J. A., Gopalan, R., and Matthews, D. R. (1991).
Publication bias in clinical research.
The Lancet, 337:867–872.

Egger, M., Smith, G. D., Schneider, M., and Minder, C. (1997a).
Bias in meta-analysis detected by a simple, graphical test.
British Medical Journal, 315:629–634.

Egger, M., Zellweger-Zähner, T., Schneider, M., Junker, C., Lengeler, C., and Antes, G. (1997b).
Language bias in randomised controlled trials published in English and German.
The Lancet, 350:326–329.

Gøtzsche, P. C. (1989).
Multiple publication of reports of drug trials.
European Journal of Clinical Pharmacology, 36:429–432.

Harbord, R. M., Egger, M., and Sterne, J. A. (2006).
A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints.
Statistics in Medicine, 25(20):3443–3457.

References III

Hedges, L. V. and Vevea, J. L. (1996).
Estimating effect size under publication bias: Small sample properties and robustness of a random effects selection model.
Journal of Educational Statistics, 21:299–332.

Higgins, J. P. and Green, S. (2009).
Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2.
Chichester.

Ioannidis, J., Munafo, M., Fusar-Poli, P., Nosek, B., and David, S. (2014).
Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention.
Trends in Cognitive Sciences, 18(5):235–241.
doi: 10.1016/j.tics.2014.02.010.

Ioannidis, J. P. A. and Trikalinos, T. A. (2007).
The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey.
Canadian Medical Association Journal, 176(8):1091–1096.

Lau, J., Ioannidis, J. P. A., Terrin, N., Schmid, C. H., and Olkin, I. (2006).
The case of the misleading funnel plot.
British Medical Journal, 333:597–600.

Light, R. J. and Pillemer, D. B. (1984).
Summing up. The science of reviewing research.
Harvard University Press, Cambridge, Massachusetts.

Macaskill, P., Walter, S. D., and Irwig, L. (2001).
A comparison of methods to detect publication bias in meta-analysis.
Statistics in Medicine, 20:641–654.

References IV

Moore, R. A., Tramer, M. R., Carroll, D., Wiffen, P. J., and McQuay, H. J. (1998).
Quantitive systematic review of topically applied non-steroidal anti-inflammatory drugs.
British Medical Journal, 316(7128):333–338.

Moreno, S., Sutton, A., Ades, A., Stanley, T., Abrams, K., Peters, J., and Cooper, N. (2009a).
Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study.
BMC Medical Research Methodology, 9:2.

Moreno, S. G., Sutton, A. J., Turner, E. H., Abrams, K. R., Cooper, N. J., Palmer, T. P., and Ades, A. E. (2009b).
Novel methods to deal with publication biases: secondary analysis of antidepressant trials in the FDA trial registry database and related journal publications.
BMJ, 339.
doi: 10.1136/bmj.b2981.

Nieminen, P., Rücker, G., Miettunen, J., Carpenter, J. R., and Schumacher, M. (2007).
Statistically significant papers in psychiatry were cited more often than others.
Journal of Clinical Epidemiology, 60(9):939–946.

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., and Rushton, L. (2006).
Comparison of two methods to detect publication bias in meta-analysis.
Journal of the American Medical Association, 295:676–680.

Rothstein, H. R., Sutton, A. J., and Borenstein, M. (2005).
Publication bias in meta analysis: prevention, assessment and adjustments.
Wiley, Chichester.

Rücker, G., Carpenter, J., and Schwarzer, G. (2011a).
Detecting and adjusting for small-study effects in meta-analysis.
Biometrical Journal, 53(2):351–368.

References V

Rücker, G., Schwarzer, G., Carpenter, J., Binder, H., and Schumacher, M. (2011b).
Treatment effect estimates adjusted for small-study effects via a limit meta-analysis.
Biostatistics, 12(1):122–142.
Doi:10.1136/jme.2008.024521.

Rücker, G., Schwarzer, G., and Carpenter, J. R. (2008).
Arcsine test for publication bias in meta-analyses with binary outcomes.
Statistics in Medicine, 27(5):746–763.

Schwarzer, G., Antes, G., and Schumacher, M. (2002).
Inflation of type I error rate in two statistical tests for the detection of publication bias in meta-analyses with binary outcomes.
Statistics in Medicine, 21:2465–2477.

Schwarzer, G., Antes, G., and Schumacher, M. (2007).
A test for publication bias in meta-analysis with sparse binary data.
Statistics in Medicine, 26:721–733.

Schwarzer, G., Carpenter, J. R., and Rücker, G. (2010).
Empirical evaluation suggests Copas selection model preferable to trim-and-fill method for selection bias in meta-analysis.
Journal of Clinical Epidemiology, 63:282–288.

Stanley, T. D. (2008).
Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection.
Oxford Bulletin of Economics and Statistics, 70:105–127.

Sterling, T. D., Rosenbaum, W. L., and Weinkam, J. J. (1995).
Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa.
The American Statistician, 49:108–112.

References VI

Sterne, J. A. C. and Egger, M. (2001).
Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis.
Journal of Clinical Epidemiology, 54:1046–1055.

Sterne, J. A. C., Gavaghan, D., and Egger, M. (2000).
Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature.
Journal of Clinical Epidemiology, 53:1119–1129.

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., and Higgins, J. P. T. (2011).
Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials.
British Medical Journal, 343:d4002.
doi: 10.1136/bmj.d4002.

Terrin, N., Schmid, C. H., and Lau, J. (2005).
In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias.
Journal of Clinical Epidemiology, 58(9):894–901.

Vevea, J. L. and Hedges, L. V. (1995).
A general linear model for estimating effect size in the presence of publication bias.
Psychometrika, 60:419–435.

Williamson, P. R. and Gamble, C. (2005).
Identification and impact of outcome selection bias in meta-analysis.
Statistics in Medicine, 24(10):1547–1561.