

the computation of maximum likelihood difference scaling (MLDS).

Classical Test Theory (CTT)

The package **psychometric** contains functions useful for correlation theory, meta-analysis (validity-generalization), reliability, item analysis, inter-rater reliability, and classical utility. Cronbach alpha, kappa coefficients, and intra-class correlation coefficients (ICC) can be found in the **psy** package.

A number of routines for scale construction and reliability analysis useful for personality and experimental psychology are contained in the packages **psych** and **MiscPsycho**. Additional measures for reliability and concordance can be computed with the **concord** package.

Other related packages

Latent class analysis can be performed using the function `lca()` from package **e1071**. Further packages are **mmlcr** and **poLCA**. They compute mixed-mode latent class regression and polytomous variable latent class analysis, respectively.

The **cfa** package allows for the computation of simple, more-sample, and stepwise configural frequency analysis (CFA).

Coefficients for interrater reliability and agreements can be computed with the **irr** package. Psychophysical data can be analyzed with the **psyphy** package. Bradley-Terry models for paired comparisons are implemented in the package **BradleyTerry** and in **eba**. The latter allows also for the computation of elimination-by-aspects models. Confidence intervals for standardized effect sizes can be found in **MBESS**.

Bibliography

J. de Leeuw and P. Mair. An Introduction to the Special Volume on Psychometrics in R *Journal of Statistical Software*, 20(1):1-5, 2007. URL <http://www.jstatsoft.org/v20/i01/>

Patrick Mair, Reinhold Hatzinger
 Department of Statistics and Mathematics
 Wirtschaftsuniversität Wien
 patrick.mair@wu-wien.ac.at
 reinhold.hatzinger@wu-wien.ac.at

meta: An R Package for Meta-Analysis

by G. Schwarzer

Introduction

The statistical method of meta-analysis can be used to combine two or more individual study results. More specifically, an overall effect is estimated by calculating a weighted average of estimates in individual studies. Various methods of meta-analysis exist that differ mainly in the weighting scheme utilised. Meta-analysis provides a statistical method to evaluate the direction and size of the effect as well as the question whether the effect is consistent across different studies.

The package **meta** is a comprehensive set of functions for meta-analysis. Initially, the package was intended to provide statistical methods for meta-analysis available in Review Manager, Version 4 (RevMan 4), the Cochrane Collaboration's program for preparing and maintaining Cochrane reviews (see <http://www.cc-ims.net/RevMan/>). The statistical capabilities of RevMan 4 have been extended over time.

The package provides methods for meta-analysis of studies comparing two groups with either binary or continuous outcome (function `metabin()`

and `metacont()`, respectively). Furthermore, the package can be used in a more general way by using the function `metagen()`, e.g., to combine hazard ratios for survival outcomes.

Statistical methods for fixed effect and random effects models (Fleiss, 1993) are available as well as functions to draw the most commonly used graphical displays (forest plots, funnel plots, and radial plots). Various statistical tests for funnel plot asymmetry, which is often taken as an indication of publication bias, are implemented. Some additional functions are available which will not be described in detail in this article, e.g. the function `trimfill()` for the trim-and-fill method, which can be used to correct for funnel plot asymmetry (Duval and Tweedie, 2000), and the function `read.mtv()` to read data analysis files exported from RevMan 4.

Another package for meta-analysis exists, called **rmeta**, which also provides functions for fixed effect and random effects models. As compared to the package **meta**, functions implementing statistical methods for funnel plot asymmetry, specific methods for continuous outcomes, the Peto method for pooling as well as the additional functions mentioned in the last paragraph are not available in the package **rmeta**.

Meta-analysis of binary outcomes

The function `metabin()` provides methods for the meta-analysis of studies comparing two groups with binary outcomes. Output from `metabin()` is an object with classes "metabin" and "meta".

The summary measure to compare the two groups (parameter `sm`) can be

- the relative risk (RR, default): $\frac{p_E}{p_C}$
- the odds ratio (OR): $\frac{p_E/(1-p_E)}{p_C/(1-p_C)}$
- the risk difference (RD): $p_E - p_C$
- the arcsine difference (AS): $\arcsin(\sqrt{p_E}) - \arcsin(\sqrt{p_C})$

with p_E and p_C denoting the event probabilities for experimental and control group, respectively.

At a minimum, the number of events (`event.e`, `event.c`) and the sample size in both groups (`n.e`, `n.c`) are needed as input to `metabin()`. Further parameters could be set, e.g., the type of fixed effect method to be utilised (parameter `method`):

- Inverse variance method, available for all effect measures (Fleiss, 1993),
- Mantel-Haenszel method, available for RR, OR, and RD (Greenland and Robins, 1985; Robins et al., 1986),
- Peto method, only available for OR (Yusuf et al., 1985).

In any case, results of a random effects meta-analysis based on the inverse variance method and a method of moments estimator for the between-study variance are also calculated (DerSimonian and Laird, 1986).

Zero cell frequencies

Several parameters of `metabin()` are concerned with the handling of studies with zero cell frequencies.

The estimated event probability \hat{p}_E or \hat{p}_C is zero if either the number of events in the experimental or control group is zero; accordingly, the estimated odds ratio and risk ratio are either 0 or infinite. If both event numbers are zero, odds ratio and relative risk are undefined (see below). Furthermore, variance estimates of log odds ratio and log relative risk are infinite due to division by zero. For the risk difference, the estimated effect is always finite, but the variance estimate can be zero. Thus, an adjustment for zero cell frequencies is necessary for odds ratio, relative risk, and risk difference. On the other hand, no adjustment for zero cell frequencies is necessary for the arcsine difference as summary measure.

It is common practice to add a small constant, typically 0.5, to each cell count in the case of zero cell frequencies (Gart and Zweifel, 1967; Pettigrew et al., 1986); a different value for this increment can be chosen (parameter `incr`). This modification can be used in general, but it is typically only applied if any of the cell counts is zero.

Three meta-analytic strategies are implemented in `metabin()` to adjust for zero cell frequencies:

- add 0.5 only to cell counts of two-by-two tables with zero cell counts (default),
- add 0.5 to all two-by-two tables in the case of zero cell counts in one or more studies (parameter `allincr=TRUE`),
- add 0.5 to all two-by-two tables (parameter `addincr=TRUE`).

To calculate the Mantel-Haenszel and the Peto estimate, there is no need to adjust for zero cell frequencies. However, an adjustment is utilised for the Mantel-Haenszel method in commonly used software for meta-analysis like RevMan 4 or the Stata procedure `metan` (<http://www.stata.com/>). Accordingly, an adjustment is also used by default for the Mantel-Haenszel method in `metabin()`; the exact Mantel-Haenszel estimate without adjustment can be calculated by setting the parameter `MH.exact=TRUE`.

For odds ratio and relative risk, studies with zero events in both groups are typically excluded from the meta-analysis, which is the default behaviour in `metabin()`. However, it is possible to include these studies in the meta-analysis by setting the parameter `allstudies=TRUE`.

Example: aspirin in myocardial infarction

A dataset of seven randomised controlled trials of the effectiveness of aspirin versus placebo in preventing death after myocardial infarction (Fleiss, 1993) is included in the package `meta`.

The result of conducting a meta-analysis of these trials using the odds ratio as measure of treatment effect is given in Figure 1. The assignment

```
> m1 <- metabin(...)
```

results in an object of classes "metabin" and "meta". Accordingly, the command

```
> m1
```

utilises the generic function `print.meta()` to print individual results of the seven trials as well as summaries for the fixed effect and random effects model. The columns `%W(fixed)` and `%W(random)` give the percentage weights of individual trials in the fixed effect and random effects model, respectively.

```

> data("Fleiss93")
> m1 <- metabin(event.e, n.e, event.c, n.c,
+               data=Fleiss93, studlab=paste(study, year),
+               sm="OR")
> m1

```

	OR	95%-CI	%W(fixed)	%W(random)
MRC-1 1974	0.7197	[0.4890; 1.0593]	3.18	8.21
CDP 1976	0.6808	[0.4574; 1.0132]	3.10	7.85
MRC-2 1979	0.8029	[0.6065; 1.0629]	5.68	13.23
GASP 1979	0.8007	[0.4863; 1.3186]	1.80	5.36
PARIS 1980	0.7981	[0.5526; 1.1529]	3.22	8.89
AMIS 1980	1.1327	[0.9347; 1.3728]	10.15	20.70
ISIS-2 1988	0.8950	[0.8294; 0.9657]	72.88	35.77

Number of trials combined: 7

	OR	95%-CI	z	p.value
Fixed effects model	0.8969	[0.8405; 0.9570]	-3.2876	0.001
Random effects model	0.8763	[0.7743; 0.9917]	-2.0918	0.0365

Quantifying heterogeneity:
 $\tau^2 = 0.0096$; $H = 1.29$ [1; 1.99]; $I^2 = 39.7\%$ [0%; 74.6%]

Test of heterogeneity:

Q	d.f.	p.value
9.95	6	0.1269

Method: Mantel-Haenszel method

Figure 1: Meta-analysis with binary outcome – myocardial infarction trials (Fleiss, 1993); output of function `metabin()`.

Summary and forest plot

The command

```
> summary(m1)
```

results in the same output as given in Figure 1 but omitting results for individual studies; actually, the function `summary.meta()` is called inside `print.meta()`. The function `summary.meta()` results in an object of class "summary.meta" with corresponding function `print.summary.meta()`.

An object of class "summary.meta" contains information on several summary statistics. For example, the command

```
> summary(m1)$fixed
```

gives a list of elements with results for the fixed effect meta-analysis.

The function `summary.meta()` can also be used to conduct a sub-group analysis. For example, the following command results in a sub-group analysis of the myocardial infarction trials based on the year of publication:

```
> summary(m1, byvar=Fleiss93$year<1980,
+          bylab="year<1980")
```

The result of a meta-analysis can be shown graphically by using the function `plot.meta()`. Figure 2 was generated by the command

```
> plot(m1, comb.f=TRUE, comb.r=TRUE)
```

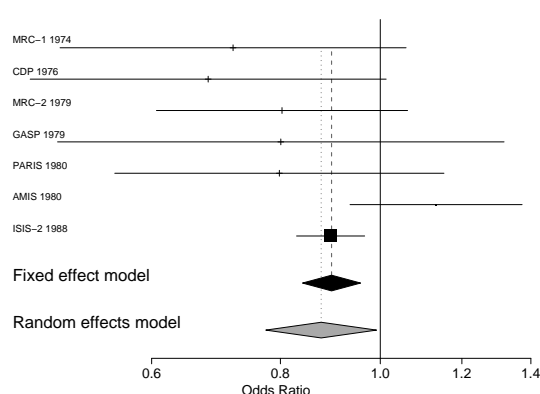


Figure 2: Forest plot of the myocardial infarction trials (Fleiss, 1993); output of function `plot.meta()`.

This type of figure is usually called a forest plot. For individual trials, the estimated odds ratio with 95% confidence interval is plotted. Fixed effect and random effects estimate and 95% confidence interval are depicted by the diamonds.

Meta-analysis of continuous outcomes

The function `metacont()` provides methods for the meta-analysis of continuous outcomes; an object of classes "metacont" and "meta" is generated. The following summary measures (parameter `sm`) are available:

- weighted mean difference (WMD, default): $\bar{x}_E - \bar{x}_C$
- standardised mean difference (SMD): $\frac{\bar{x}_E - \bar{x}_C}{SD}$

with \bar{x}_E and \bar{x}_C denoting the mean values in the two groups and SD denoting the average standard deviation. Hedges' adjusted g (Cooper and Hedges, 1994) is utilised for the standardised mean difference.

At a minimum, sample sizes (`n.e`, `n.c`), mean values (`mean.e`, `mean.c`), and standard deviations (`sd.e`, `sd.c`) are needed as input to `metacont()`. Both fixed effect and random effects summary estimates are calculated based on the inverse variance method (Fleiss, 1993).

The generic functions `print()`, `summary()`, and `plot()` described in the last section are also available for meta-analyses with continuous outcomes.

Meta-analysis based on generic inverse variance method

The function `metagen()` provides methods for the meta-analysis of any outcome. At a minimum, the estimated effect (TE) and its standard error (`seTE`) are needed as input to `metagen()`. Both fixed effect and random effects summary estimates are calculated based on the inverse variance method (Fleiss, 1993).

For example, the function `metagen()` can be utilised to summarise

- adjusted estimates (e.g. from logistic regression),
- log hazard ratios for survival outcomes (Parmar et al., 1998),
- estimates from cross-over trials (Curtin et al., 2002),
- estimates from both direct and indirect comparisons (Bucher et al., 1997).

The generic functions `print()`, `summary()`, and `plot()` described in the section on meta-analyses of binary outcomes can be utilised for objects generated by `metagen()`.

Statistical methods to detect small-study effects

In meta-analyses it sometimes happens that smaller studies show different, often larger, treatment effects. One possible reason for such "small study effects" is publication bias. This is said to occur when the chance of a smaller study being published is increased if it shows a stronger effect. If this occurs, it in turn biases the result of the meta-analysis. A comprehensive review of these issues is given in Rothstein et al. (2005).

Funnel plot

A natural way of visualising the evidence for possible small study effects/publication bias is the funnel plot (Light and Pillemer, 1984), which plots each study's treatment effect (x-axis) against a measure of its variability (y-axis); usually this is the standard error, although other options are preferable in different situations (Sterne and Egger, 2001). The funnel plot gives an idea of whether there is any dependence of treatment effect on precision.

In principal, at least two sources of asymmetry in funnel plots exist. Publication bias, the first of them, is well known. The second reason is heterogeneity, for example, smaller studies may select patients who are more likely to benefit from the intervention. Effects like these have been referred to as "small study effects" (Sterne et al., 2000).

A funnel plot of the myocardial infarction trials generated by the command

```
> funnel(m1, level=0.95)
```

is plotted in Figure 3. In this figure, a gap in the lower right part is somewhat suggestive of asymmetry in the funnel plot.

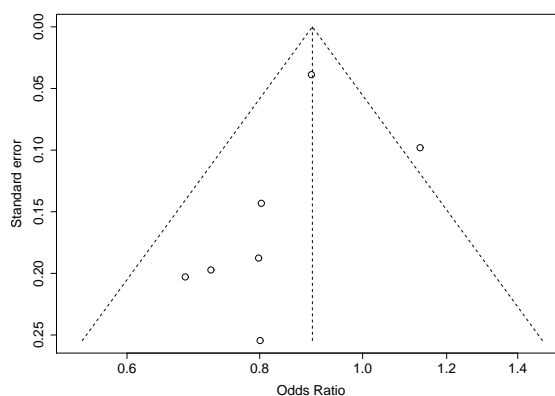


Figure 3: Funnel plot of the myocardial infarction trials (Fleiss, 1993); output of function `funnel()`.

Tests for funnel plot asymmetry

A number of tests for small study effects/publication bias have been developed and are available in the function `metabias()`. Output of this command is an object of class "htest" which utilises the default `print()` function available with basic R.

Tests for small study effects/publication bias fall into two classes:

- non-parametric tests using rank-correlation methods, going back to [Begg and Mazumdar \(1994\)](#),
- regression tests, represented by the so-called Egger test ([Egger et al., 1997](#)).

The tests assume that under the null hypothesis of no "publication bias" among studies included in a meta-analysis, there is no association between treatment effect and precision.

While this assumption is plausible when the outcome is quantitative, as assuming normality the sample mean is statistically independent of the sample variance, it is not generally true for binary data. Specifically, suppose the outcome is binary and the effect is summarised by the log relative risk (logRR) or log odds ratio (logOR). The variance estimators of both the logRR and logOR are statistically dependent on the estimated logRR and logOR. Even in the absence of small study effects, this dependence induces asymmetry in the funnel plot ([Macaskill et al., 2001](#); [Schwarzer et al., 2002](#)).

This observation has motivated recent proposals to modify existing tests for binary outcomes ([Harbord et al., 2006](#); [Peters et al., 2006](#); [Schwarzer et al., 2007](#); [Rücker et al., 2007](#)). These tests are available in the function `metabias()` (parameter `method`).

In [Figure 4](#), results for two different tests on funnel plot asymmetry ([Begg and Mazumdar, 1994](#); [Harbord et al., 2006](#)) are given for the meta-analysis of the myocardial infarction trials. Both tests are non-significant ($p\text{-value} \geq 0.1$); thus, there is no clear indication of funnel plot asymmetry.

Summary

The package `meta` is a comprehensive set of functions for meta-analysis and provides functions for the most commonly used outcomes, i.e., `metabin()` for binary and `metacont()` for continuous outcomes. Furthermore, the package can be used in a more general way by using the function `metagen()`. Generic functions to print and plot the results of a meta-analysis and to conduct sub-group analyses are available. Various statistical tests for funnel plot asymmetry which is often taken as an indication of publication bias are implemented.

Acknowledgments

This research was supported by the Deutsche Forschungsgemeinschaft (German Research Foundation) under grants Schu 756/6-1 and FOR 534.

Bibliography

- C. B. Begg and M. Mazumdar. Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50:1088–1101, 1994.
- H. C. Bucher, G. H. Guyatt, L. E. Griffith, and S. D. Walter. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology*, 50: 683–91, 1997.
- H. Cooper and L. V. Hedges, editors. *The Handbook of Research Synthesis*. Russell Sage Foundation, Newbury Park, CA, 1994.
- F. Curtin, D. G. Altman, and D. Elbourne. Meta-analysis combining parallel and cross-over clinical trials. I: Continuous outcomes. *Statistics in Medicine*, 21:2131–2144, 2002.
- R. DerSimonian and N. Laird. Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7:177–188, 1986.
- S. Duval and R. Tweedie. A nonparametric "Trim and Fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, 95:89–98, 2000.
- M. Egger, G. D. Smith, M. Schneider, and C. Minder. Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315:629–634, 1997.
- J. L. Fleiss. The statistical basis of meta-analysis. *Statistical Methods in Medical Research*, 2:121–145, 1993.
- J. J. Gart and J. R. Zweifel. On the bias of various estimators of the logit and its variance with application to quantal bioassay. *Biometrika*, 54:181–187, 1967.
- S. Greenland and J. M. Robins. Estimation of a common effect parameter from sparse follow-up data. *Biometrics*, 41:55–68, 1985. (C/R: V45 p1323–1324).
- R. M. Harbord, M. Egger, and J. A. Sterne. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Statistics in Medicine*, 25(20):3443–3457, 2006.
- R. J. Light and D. B. Pillemer. *Summing up. The science of reviewing research*. Harvard University Press, Cambridge, Massachusetts, 1984.
- P. Macaskill, S. D. Walter, and L. Irwig. A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine*, 20:641–654, 2001.

```

> metabias(m1, meth="rank")

Rank correlation test of funnel plot asymmetry

data:  m1
z = -1.3517, p-value = 0.1765
alternative hypothesis: true is 0
sample estimates:
      ks      se.ks
-9.000000  6.658328

> metabias(m1, meth="score")

Linear regression test of funnel plot asymmetry (efficient score)

data:  m1
t = -0.9214, df = 5, p-value = 0.3991
alternative hypothesis: true is 0
sample estimates:
      bias      se.bias      slope
-0.72587833  0.78775820 -0.05932016

```

Figure 4: Statistical tests for funnel plot asymmetry – myocardial infarction trials (Fleiss, 1993); output of function metabias().

- M. K. B. Parmar, V. Torri, and L. Stewart. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Statistics in Medicine*, 17:2815–2834, 1998.
- J. L. Peters, A. J. Sutton, D. R. Jones, K. R. Abrams, and L. Rushton. Comparison of two methods to detect publication bias in meta-analysis. *Journal of the American Medical Association*, 295:676–680, 2006.
- H. M. Pettigrew, J. J. Gart, and D. G. Thomas. The bias and higher cumulants of the logarithm of a binomial variate. *Biometrika*, 73:425–435, 1986.
- J. Robins, N. E. Breslow, and S. Greenland. Estimators of the Mantel-Haenszel variance consistent in both sparse data and large-strata limiting models. *Biometrics*, 42:311–323, 1986.
- H. R. Rothstein, A. J. Sutton, and M. Borenstein. *Publication bias in meta analysis: prevention, assessment and adjustments*. Wiley, Chichester, 2005.
- G. Rücker, G. Schwarzer, and J. R. Carpenter. Arcsine test for publication bias in meta-analyses with binary outcomes. *Statistics in Medicine*, 2007. Accepted.
- G. Schwarzer, G. Antes, and M. Schumacher. 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, 2002.
- G. Schwarzer, G. Antes, and M. Schumacher. A test for publication bias in meta-analysis with sparse binary data. *Statistics in Medicine*, 26(4):721–733, 2007.
- J. A. C. Sterne and M. Egger. Funnel plots for detecting bias in meta-analysis: Guideline on choice of axis. *Journal of Clinical Epidemiology*, 54:1046–1055, 2001.
- J. A. C. Sterne, D. Gavaghan, and M. Egger. Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology*, 53:1119–1129, 2000.
- S. Yusuf, R. Peto, J. Lewis, R. Collins, and P. Sleight. Beta blockade during and after myocardial infarction: An overview of the randomized trials. *Progress in Cardiovascular Diseases*, 27:335–371, 1985.

Guido Schwarzer
 Institute of Medical Biometry and Medical Informatics
 University Medical Center Freiburg, Germany
 sc@imbi.uni-freiburg.de