copas: An R package for Fitting the Copas Selection Model

by J. Carpenter, G. Rücker and G. Schwarzer

Abstract: This article describes the R package **copas** which is an add-on package to the R package **meta**. The R package **copas** can be used to fit the Copas selection model to adjust for bias in meta-analysis. A clinical example is used to illustrate fitting and interpreting the Copas selection model.

Introduction

Systematic reviews play a key role in the evaluation of new interventions. While combining evidence from a number of studies should reduce both bias and uncertainty, this is sometimes not the case because published studies represent a biased selection of the evidence. This can happen for a number of reasons, for example authors may be more likely to submit trials with 'significant' results for publication or journals may be more likely to publish smaller trials if they have 'significant' results. Empirical studies have established evidence for these kinds of biases and others (Rothstein et al., 2005; Nieminen et al., 2007).

In consequence, considerable effort has been directed at the problem of developing reliable tests for selection biases (Harbord et al., 2006; Rücker et al., 2008), and in a second step, correcting estimates for publication bias (Rothstein et al., 2005). One of the most promising methods to date has been the so-called Copas selection model (Copas, 1999; Copas and Shi, 2000, 2001), which is derived from the Heckman 2-stage regression model (see Little and Rubin (2002)).

Comprehensive evaluations suggest that this approach (i) can provide a useful summary in around 80% of meta-analyses (Carpenter et al., 2009) and (ii) is preferable to the trim-and-fill method to adjust for bias in meta-analysis (Schwarzer et al., 2009). This article describes the R package **copas** for fitting and interpreting the Copas selection model.

Copas selection model

We first briefly describe the model. The Copas selection model has two components: (i) a model for the outcome, measured on a chosen scale, e.g. the log odds ratio, log risk ratio or the arcsine difference (Rücker et al., 2009), and (ii) a 'selection' model giving the probability that study i is observed/published. A correlation parameter ρ between these two components models the extent of selection/publication bias; the stronger the correlation, the greater the chance that only more extreme outcomes are observed/published.

In more detail, let (ϵ_i, δ_i) follow a bivariate normal distribution with mean 0 and covariance matrix

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Denote the underlying population intervention effect by θ, and between study heterogeneity variance τ^2. For each study i in a meta-analysis, let $\hat{\theta}_i$ denote the estimate of θ and s_i^2 the estimated variance of $\hat{\theta}_i$, whose true, unobserved, variance is σ_i^2. We model the observed outcome of study i using the usual random effects model, i.e. as

$$\hat{\theta}_i = \theta + \sqrt{(\tau^2 + \sigma_i^2)}\epsilon_i \tag{1}$$

and say study i is observed/published if $Z_i > 0$, where

$$Z_i = \gamma_0 + \gamma_1/s_i + \delta_i \tag{2}$$

with fixed γ_0 and γ_1. From (2), the marginal probability that study i is observed is

$$\Pr(Z_i > 0) = \Pr(\delta_i > -\gamma_0 - \gamma_1/s_i) \tag{3}$$
$$= \Phi(\gamma_0 + \gamma_1/s_i), \tag{4}$$

where $\Phi(.)$ is the cumulative density function of the standard normal. Thus $\Phi(\gamma_0)$ can be interpreted as the marginal probability of publishing a study with infinite standard error, and γ_1 is associated with the change in publication probability with increasing precision. Note that the appearance of s_i in (2) means that the probability of publication reflects the sampling variance of study i.

Copas (1999) and Copas and Shi (2000) use standard properties of the normal distribution to show that the probability of observing study i is

$$\Phi\left\{ \frac{\gamma_0 + \gamma_1/s_i + \rho\sigma_i(\hat{\theta}_i - \theta)/(\sigma_i^2 + \tau^2)}{\sqrt{1 - \rho^2\sigma_i^2/(\sigma_i^2 + \tau^2)}} \right\}. \tag{5}$$

Thus if $\rho = 0$, (1) and (2) are unrelated and a meta-analysis of observed studies will give an approximately unbiased estimate of θ. Conversely, if large θ means a strong treatment effect and $\rho > 0$, then the probability of observing study i is increased the larger $\hat{\theta}_i$. In this situation, a meta-analysis of observed studies will give a biased estimate of θ.

Fitting the Copas selection model

We have developed the R package **copas** to provide a comprehensive set of R functions for fitting the

```
> ## Read in data
> data(Crowther2003)
> ## Do meta-analysis using function metabin() from R package meta
> m.crowther <- metabin(event.e, total.e, event.c, total.c,
+                       data=Crowther2003, sm="OR", studlab=study)
Warning message:
In metabin(event.e, total.e, event.c, total.c, data = Crowther2003,  :
  Increment 0.5 added to each cell in 2x2 tables with zero cell frequencies
> ## Do test for funnel plot asymmetry (Harbord et al.)
> ## using function metabias() from R package meta
> metabias(m.crowther, meth="score")

Linear regression test of funnel plot asymmetry (efficient score)

data:  m.crowther
t = -3.4551, df = 7, p-value = 0.01062
alternative hypothesis: asymmetry in funnel plot
sample estimates:
     bias    se.bias      slope
-2.6149672  0.7568465  0.2927025


> ## Do Copas analysis
> cop1 <- copas(m.crowther)
> ## Plot Copas analysis
> plot(cop1)
> ## Redo Copas analysis as P-value of residual selection bias
> ## is still significant
> cop2 <- copas(m.crowther,
+              gamma0.range=c(-0.55, 2),
+              gamma1.range=cop1$gamma1.range)
Warning messages:
1: In sqrt(1 - rho.tilde^2) : NaNs produced
2: In sqrt((tau^2 + sigma^2) * (1 - rho.tilde^2)) : NaNs produced
3: In sqrt((1 - rho.tilde^2)) : NaNs produced
> ## Plot Copas analysis
> plot(cop2)
> ## Print summary of Copas analysis
> summary(cop2)
Summary of Copas selection model analysis:

              publprob     OR            95%-CI pval.treat  pval.rsb N.unpubl
                  1.00 0.4967  [0.3247; 0.7599]     0.0013     0.006        0
                  0.82 0.5483  [0.3494; 0.8605]      0.009     0.007        1
                  0.67 0.6063  [0.3938; 0.9335]      0.023    0.0115        2
                  0.55 0.6702  [0.4601; 0.9761]      0.037    0.0205        4
                  0.45 0.7402  [0.5376; 1.0193]     0.0653     0.046        6
                  0.37 0.8337  [0.6055; 1.1480]     0.2652    0.2461        9


   Copas model (adj) 0.8337  [0.6055; 1.1480]     0.2652    0.2461        9
 Random effects model 0.4880  [0.3234; 0.7363]     0.0006


 Legend:
 publprob  - Probability of publishing the study with the largest standard error
 pval.treat - P-value for hypothesis that the treatment effect is equal in both groups
 pval.rsb  - P-value for hypothesis that no further selection remains unexplained
 N.unpubl  - Approximate number of studies the model suggests remain unpublished
```

Figure 1: Example of Copas analysis of phenobarbital versus control for reducing neonatal periventricular haemorrhage (Crowther and Henderson-Smart, 2003); output of function summary.copas().

Copas selection model, then printing and displaying the results. The R package **copas** is an add-on package to the R package **meta** (Schwarzer, 2007). To illustrate the use of the R package **copas**, consider a meta-analysis of 9 studies comparing prophylactic maternal phenobarbital with control for periventricular haemorrhage in preterm infants (Crowther and Henderson-Smart, 2003).

Figure 1 illustrates the analysis of these data, which have a binary response (presence or absence of haemorrhage). After reading the data in, we perform a meta-analysis on the odds-ratio scale using function `metabin()` from the R package **meta**. Note the warning that a continuity correction has been used, since one study has a zero cell. By default, 0.5 is added only to cell counts of two-by-two tables with zero cell frequencies. Two other stategies to adjust for zero cell frequencies are implemented in the function `metabin()` (Schwarzer, 2007): (i) add 0.5 to all two-by-two tables in the case of zero cell counts in one or more studies (`allincr = TRUE`), (ii) add 0.5 to all two-by-two tables irrespective of zero cell counts (`addincr = TRUE`).

The random effects model suggests a significant intervention benefit (p = 0.0006). Nevertheless, a funnel plot of the data (top left panel Figure 2) suggests effects may be systematically larger in smaller studies. Investigating this with a statistical test (function `metabias()`, Figure 1) supports this suspicion with $p = 0.011$.

We therefore perform a Copas analysis, using the function `copas()`, as shown in Figure 1. This fits the Copas selection model repeatedly, by maximising the likelihood (Carpenter et al., 2009) — subject to the constraints that $-1 < \rho < 1$ and $\tau^2 \geq 0$ — over a grid of (γ_0, γ_1) values using the existing R function `optim()` (L-BFGS-B method). We use transformations of ρ and τ^2 in the log-likelihood to reduce numerical instability, which is a known issue with this model (Carpenter et al., 2009). We redo the Copas analysis extending the range of values for γ_0 as the P-value of residual selection bias is still significant using the default settings. Three warning messages are printed concerning a parameter `rho.tilde` which is used internally in likelihood estimation (Copas and Shi, 2000, p. 250). Typically, these warnings can be safely ignored. The function `plot.copas()` can be used as a diagnostic tool (see Discussion).

An object of class `copas` is created which we can `print`, `summarise` and `plot`. Figure 2 shows the result of function `plot.copas()`, which by default displays four panels. The top left plot is the usual funnel plot, which plots the study specific effect sizes (here log odds ratios) against their standard error. The vertical broken line is the usual fixed effect estimate of the treatment effect, while the vertical grey line is the usual random effects estimate of the treatment ef-

fect. Diagonal broken lines show ± 2 standard errors about the fixed effect. If there is no heterogeneity or selection bias, we would expect about 95% of studies to lie in this 'funnel'. In this example, there is a suggestion that smaller studies tend to show a stronger effect.

Given the range of study standard errors in the meta-analysis, the function `copas()` chooses a range of γ_0 and γ_1 values for the selection model (2). These are chosen to represent varying selection strength. Specifically, the probability of publishing the study with the largest SE (often the smallest study) ranges from around 0.3 to 1. In this example, the initial analysis (Figure 1, with the default ranges for (γ_0, γ_1)) has the lower bound for γ_0 at -0.45, which just stops short of the region where the degree of selection is sufficient to explain the asymmetry in the funnel plot. We thus repeat the analysis with a slightly larger range for γ_0, and it is this second analysis that is shown in Figure 2.

The function `copas()` fits the Copas selection model over a grid of (by default 400) points. The top right panel produced by `plot.copas` shows a contour plot of the resulting treatment estimates $(\hat{\theta})$ over the range of (γ_0, γ_1) values. Contours are labeled with values of $\hat{\theta}$, (which can be specified by the user) in this case $-0.6, -0.5, -0.4, \dots$. The contour plot suggests that as selection increases (i.e. as we move away from the top right) the treatment estimate declines, but the contours are locally parallel. As described in more detail in the Appendix of Carpenter et al. (2009), in such cases the contour plot can be summarised by looking at how the treatment estimate varies as selection increases along a line orthogonal to the contours. Using an algorithm we developed (Carpenter et al., 2009), this orthogonal line is estimated and superimposed on the contour plot. The places where it intersects with the contours are marked with an 'o'.

The lower two panels of Figure 2 use this information to present an accessible summary of the Copas selection model analysis. First, at each of the line/contour intersections marked with an 'o' in the contour plot, the program calculates the probability of publishing the trial with the largest SE. Then, in the lower left panel this is plotted against the corresponding treatment estimate ($\pm$ its 95% confidence interval). In our example this shows that with little or no selection, the treatment estimate is close to that obtained by the usual random effects meta-analysis.[1] As selection increases (so studies with large SE's but smaller effect sizes are less likely to make it into the meta-analysis) so the treatment estimate moves towards the null value and the significance (indicated by the degree that the 95% confidence interval overlaps 0) decreases.

Finally, the bottom right panel attempts to answer

---

[1]Exact agreement is not to be expected, as the usual random effects analysis uses a method of moments estimate of heterogeneity, whereas the Copas selection model uses a maximum likelihood estimate.
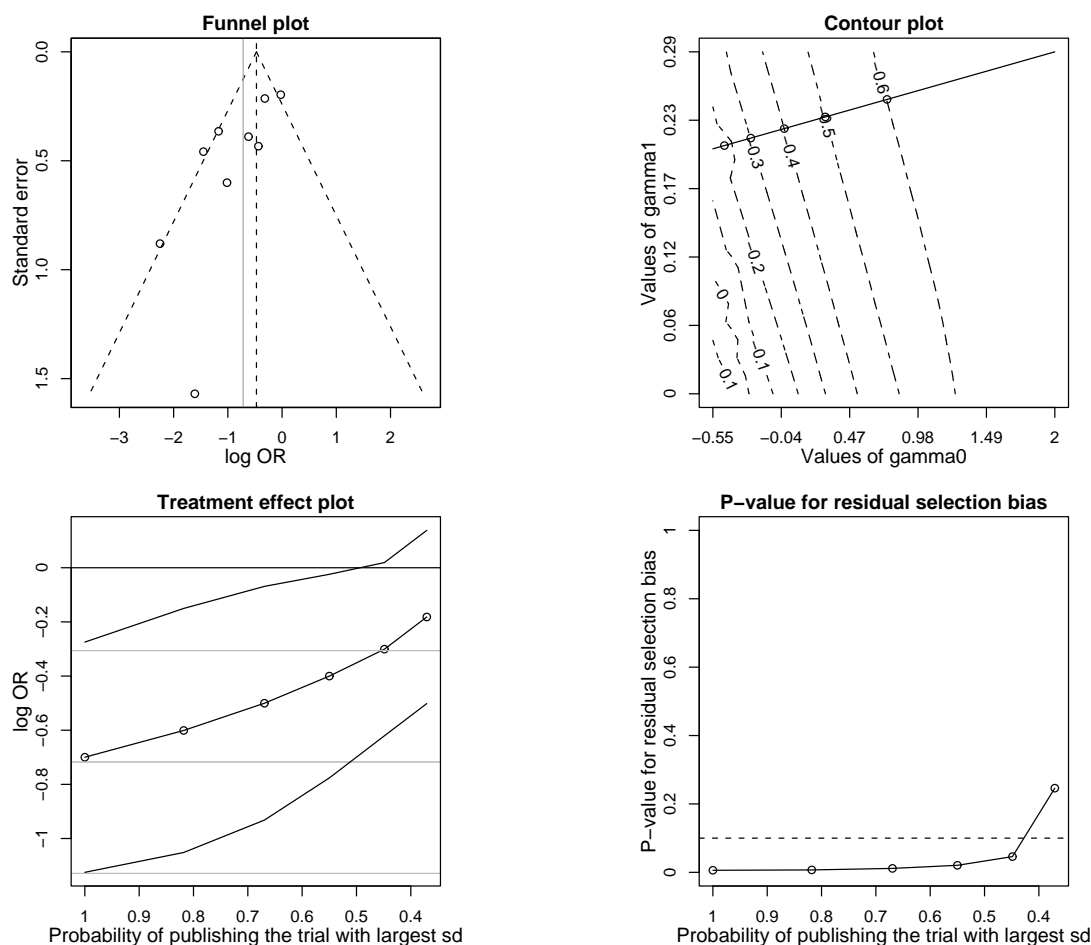
Figure 2: Plot of the results of a `copas` analysis of phenobarbital versus control for reducing neonatal periventricular haemorrhage (Crowther and Henderson-Smart, 2003).

the question of which degree of selection is most plausible under the model; i.e. which treatment estimates should we pay most attention to. It does this as follows. For each degree of selection (summarised by the probability of publishing the trial with the largest SE) the program calculates whether the remaining asymmetry in the funnel plot is more than might be expected by chance, and calculates a p-value to test this hypothesis. These are then plotted against the corresponding degree of selection. Degrees of selection corresponding to p-values above 0.1 (i.e. at this degree of selection, no evidence of residual selection bias in the data) are more plausible under the model; corresponding treatment estimates in the bottom left panel are more plausible.

In this example, with no selection bias (left hand end of bottom left panel) the p-value for residual selection bias in the data is < 0.1, suggesting the meta-analysis should not be interpreted at face value. Rather, it is only when the probability of publishing the trial with the largest SD is as low as 0.4 that the asymmetry seen in the funnel plot is explained. In

this case, the bottom left panel indicates the treatment effect is no longer significant at the 5% level. The function `summary.copas()` (Figure 1) complements this with a numerical summary.

The Copas selection model analysis therefore suggests that after accounting for selection bias and/or other small study effects, there is no evidence of a benefit of the intervention. This agrees with the authors of the original study, who comment that the two largest trials — which are of higher quality — show no benefit, and conclude "evidence does not support phenobarbital treatment to women giving birth before 34 weeks to decrease the risk of bleeding into the babies' brains".

Arguments of function `copas()`

Although the majority of studies can be analysed automatically using the Copas selection model this is not always true. Some analyses will require fine tuning. The following options are available:

- `gamma0.range`, `gamma1.range`:
  These allow users to control the grid of (γ_0, γ_1) values in the selection part of the Copas selection model (equation 2) which the program ranges over to produce the contour plot (top right panel Figure 2).

- `ngrid`:
  This parameter controls how fine the grid of (γ_0, γ_1) is. The `copas` function fits the Copas selection model over a grid of `ngrid×ngrid` values.

- `levels`:
  Fitting the Copas model over the grid specified by the arguments `gamma0.range`, `gamma1.range` and `ngrid` results in a treatment estimate at every point in the grid. These are then displayed on a contour plot where contours of treatment effect (z-axis) are shown by gamma0 (x-axis) and gamma1 (y-axis). This argument is a numeric vector which specifies the treatment effects for which contour lines will be drawn.

- `left`:
  A logical indicating whether the cause of any selection bias is due to missing studies on the left or right of the funnel plot: left hand side if `left = TRUE`, right hand side if `left = FALSE`. This information is needed in order to be sure the test for presence of residual selection bias is calculated correctly.

- `rho.bound`:
  A number giving the upper bound for the correlation parameter ρ in the Copas selection model. The default is 0.9999. At this value, warnings are sometimes triggered by the program attempting to take the square root of numbers that are just negative; in all analyses we have carried out these can safely be ignored. Alternatively, repeat the analysis with a slightly smaller bound. Values less than 0.95 are likely to cause irregularites in regions of the contour plot where there is a high degree of selection.

- `silent`:
  A logical indicating whether information on progress in fitting the Copas selection model should be printed: `silent = TRUE` specifies not to print information (the default).

- `warn`:
  A number setting the handling of warning messages. It is not uncommon for numerical problems to be encountered during estimation over the grid of (γ_0, γ_1) values. Usually this does not indicate a serious problem. This option specifies what to do with warning messages — `warn = -1`: ignore all warnings; `warn`

= 0 (the default): store warnings till the function finishes; if there are less than 10, print them, otherwise print a message saying warning messages were generated; `warn = 1`: print warnings as they occur; `warn = 2`: stop the function when the first warning is generated.

All the information used to generate the plots is available as attributes of the object created by the `copas` function. Thus tailor-made versions of the panels in Figure 2 can be created by users without any further calculation.

Finally, the `summary` function and the `plot` function allow the user to specify the confidence level.

Discussion

The extensive literature on selection bias in meta-analysis (see Rothstein et al. (2005) and references therein) reflects the importance to the community of systematic reviewers of detecting, and where possible adjusting for, selection bias. The Copas selection model is a key tool for doing this (Carpenter et al., 2009) (alongside other methods such as the trim-and-fill method included in the **meta** package (Schwarzer, 2007)).

An empirical evaluation of 157 meta-analyses with 4 to 66 studies showed that our implementation of the Copas selection model provided a useful summary in about 80% of meta-analyses (Carpenter et al., 2009). In the remaining meta-analyses (i) the contour plot did not show roughly parallel contour lines, (ii) the 95% confidence intervals in the treatment effect plot did not vary smoothly, or (iii) *P*-values in the *P*-value plot for residual selection bias were erratic. A contour plot without roughly parallel contour lines did appear in situations with an apparently symmetric funnel plot, i.e. when there was no indication of selection bias. This is not a weakness of the model or the software but a consequence of the flat likelihood and the treatment effect being invariant in this situation. Irregularities in the treatment effect plot and *P*-value plot are typically due to estimation problems. In general, problems in the estimation process can be judged by looking at the output from function `plot.copas()` which should be used routinely as a diagnostic tool.

We have attempted to make the help files for the **copas** package accessible to systematic reviewers, who in most cases are likely to be new users of R. With this package, we therefore believe that R has a powerful toolkit for systematic reviewers.

Acknowledgments

Bibliography

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

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

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

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

C. A. Crowther and D. J. Henderson-Smart. Phenobarbital prior to preterm birth for preventing neonatal periventricular haemorrhage. *Cochrane Database of Systematic Reviews*, 3, 2003. Art. No.: CD000164, doi: 10.1002/14651858.CD000164.

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. A. Little and D. B. Rubin. *Statistical Analysis with Missing Data*. John Wiley & Sons, Chichester, second edition, 2002.

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

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, J. Carpenter, and I. Olkin. Why add anything to nothing? The arcsine difference as a measure of treatment effect in meta-analysis with zero cells. *Statistics in Medicine*, 28 (5):721–738, 2009.

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

G. Schwarzer. meta: An R package for meta-analysis. *R News*, 7(3):40–45, 2007. URL http://cran.r-project.org/doc/Rnews/Rnews_2007-3.pdf.

G. Schwarzer, J. R. Carpenter, and G. Rücker. Empirical evaluation suggests Copas selection model preferable to Trim-and-Fill for selection bias in meta-analysis. *Journal of Clinical Epidemiology*, in press, doi: 10.1016/j.jclinepi.2009.05.008.

James Carpenter
Institute of Medical Biometry and Medical Informatics
University Medical Center Freiburg, Germany
jrc@imbi.uni-freiburg.de

Gerta Rücker
Institute of Medical Biometry and Medical Informatics
University Medical Center Freiburg, Germany
ruecker@imbi.uni-freiburg.de

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