metaplus: An R Package for the Analysis of Robust Meta-Analysis and Meta-Regression

by Ken J. Beath

Abstract The **metaplus** package is described with examples of its use for fitting meta-analysis and meta-regression. For either meta-analysis or meta-regression it is possible to fit one of three models: standard normal random effect, *t*-distribution random effect or mixture of normal random effect. The latter two models allow for robustness by allowing for a random effect distribution with heavier tails than the normal distribution, and for both robust models the presence of outliers may be tested using the parametric bootstrap. For the robust mixture model the outlier studies may be identified through their posterior probability of membership in the outlier component of the mixture. Plots allow the results of the different models to be compared. The package is demonstrated on three examples: a meta-analysis with no outliers, a meta-analysis with an outlier and a meta-regression with an outlier.

1 Introduction

Meta-analysis is a method of combining the results of different studies to produce one overall result. Meta-regression is an extension to meta-analysis which allows for covariates describing each study to improve the accuracy of the meta-analysis. For example we might have studies comparing a drug to placebo, with varying doses of the drug used in the different studies. It would be expected that if the drug works then it is likely, although not certain, that the effectiveness would increase with dose. By including this in the model, the unexplained variation would then be reduced.

The main difficulty with combining studies is that the differences between studies may be greater than would be indicated by the variation within each study. This was allowed for by introducing the random effect model where the effect for each study has two components: an overall effect and a random component specific to each study, with the random component assumed to have a normal distribution. The model without a random effect is known as the fixed effect model, which is equivalent to a random effect with zero variance for the random effect.

One difficulty is that the assumption of a normally distributed random effect may be unrealistic, with a common violation that the tails are heavier than would be expected. One solution to this is to use an alternative to the normal distribution for the random effect, for example the *t*-distribution, as described in Lee and Thompson (2008) and Baker and Jackson (2008). This however does not identify which studies are unusual, so a method described by Gumedze and Jackson (2011) is to assume that studies are either normal or are outlier studies from a random effect distribution with a higher variance. Only one study is assumed to be an outlier, with each study tested in turn, but multiple outliers then allowed for using order statitistics. Beath (2014) noted the similarity of this model to a mixture model, which also allowed for a more general fitting algorithm and a statistical test for the presence of outliers and indication of which studies are outliers.

The purpose of the **metaplus** package is to fit these specific models. It is not designed to replace a more general meta-analysis package, such as **metafor** (Viechtbauer, 2010) but to provide additional specialised analyses. In producing forest plots, it builds upon the functionality of **metafor**, allowing the various models to be compared.

 ISSN 2073-4859

1.1 Models

The random effect meta-analysis model assumes that the observed treatment effect Y_i for study i is

$$Y_i = \mu + E_i + \epsilon_i,$$

where μ is the overall mean for the studies, E_i is a random effect with mean zero, and ϵ_i is a normally distributed error with variance σ_i^2 for study i, where the within study variance σ_i^2 is assumed to be known.

An extension to the random effect meta-analysis model is to include covariates to explain (Berkey et al., 1995) the heterogeneity. Incorporating this into the meta-analysis model we obtain:

$$Y_i = \mu + X_i^T \beta + E_i + \epsilon_i,$$

where X_i is a vector of covariate values for study i, and β is a vector of the corresponding parameters.

In **metaplus** there are three available random effect distributions:

Normal The probability density function for study i is

$$f\left(Y_i|X_i;\mu,\tau\right) = \frac{1}{\sqrt{2\pi\left(\sigma_i^2+\tau^2\right)}} \exp\left(-\frac{\left(Y_i-\mu-X_i^T\beta\right)^2}{2\left(\sigma_i^2+\tau^2\right)}\right).$$

Robust t-distribution Introduced as one of a number of distributions for robust meta-analysis by Lee and Thompson (2008) and Baker and Jackson (2008). This replaces the normal random effect distribution with a t-distribution. The degrees of freedom (ν) of the t-distribution controls the heaviness of the tails, and is estimated from the data. The probability density function no longer has a closed-form expression, requiring integration over the t-distribution random effect as

$$f\left(Y_i|X_i;\mu,\tau,\nu\right) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \int_{-\infty}^{\infty} \exp\left(\frac{-\left(Y_i-\mu-X_i^T\beta-\eta\right)^2}{2\sigma_i^2}\right) g\left(\eta|\tau,\nu\right) d\eta,$$

where $g\left(\eta|\tau,\nu\right)$ is the distribution function of a t-distribution with ν degrees of freedom

$$g\left(\eta|\tau,\nu\right) = \frac{\Gamma\left(\left(\nu+1\right)/2\right)}{\tau\sqrt{\pi\nu}\Gamma\left(\nu/2\right)} \left(1+\frac{\eta^2}{\nu\tau^2}\right)^{-\left(\left(\nu+1\right)/2\right)}.$$

Robust mixture This assumes that a study can belong to one of two classes, with each class a standard random effect model with the same mean but different random effect variance, which is higher for the outlier class (Beath, 2014). The robust meta-analysis model takes the form

$$Y_{i|k} = \mu + E_{i|k} + \epsilon_i,$$

where ϵ_i is as for the standard model, but $E_{i|k}$ is now a random effect dependent on the class, where $k = 1, 2$ indexes the classes, with $k = 1$ corresponding to standard studies and $k = 2$ to outlier studies, with random effect variances τ_1^2, τ_2^2 respectively, with the restriction that $\tau_2^2 > \tau_1^2$, and again zero mean. The probability density function becomes the weighted sum of the probability density function for each class, with weights the proportion of studies in each class π_1, π_2 for the standard and outlier studies, respectively, as

$$f\left(Y_i|X_i;\mu,\tau_1,\tau_2,\pi_1,\pi_2\right)=\sum_{k=1}^{2}\pi_k\frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sigma_i^2+\tau_k^2}\right)^{1/2}\exp\left(-\frac{1}{2}\frac{\left(Y_i-\mu-X_i^T\beta\right)^2}{\sigma_i^2+\tau_k^2}\right).$$

with the constraint that $\pi_1+\pi_2=1$

1.2 Profile likelihood confidence intervals

A difficulty with the use of standard maximum likelihood techniques for random effect models is that it produces biased estimates for the variance of the random effect, which results in biased estimates of the standard errors for the parameters of interest, and therefore poor coverage. The solution for meta-analysis has been the use of REML, however this is difficult for the robust models. However, profile likelihood based confidence intervals (Pawitan, 2001)[p 61] have been found to be superior (Hardy and Thompson, 1996) to those obtained using REML, and they are used for all fitted models. The profile likelihood confidence intervals are obtained from routines based on the `mle2` procedure in the package **bbmle** which provides an extended version of `mle`.

1.3 Parametric boostrap

Testing for the need for the robust distributions requires a test of $\nu=0$ for the t-distribution and $\pi_2=0$ for the robust mixture. Both tests involve a test of a parameter on the edge of the parameter space, so the usual asymptotic theory cannot be used. The solution is to use the parametric bootstrap, which involves simulating data sets under the null hypothesis and calculating the test statistic for each. The observed test statistic is then compared to the simulated test statistics to calculate a p-value. Application of the parametric bootstrap to mixture models is described in McLachlan (1987).

1.4 Other computational details

For both the robust models the starting values are important, as the optimisation may converge to a local minima. For the t-distribution, starting values are obtained by first fitting a standard normal random effect model, and the starting values from this model used with a range of starting values for the t-distribution degrees of freedom, and the maximum likelihood chosen. Numerical integration was used to obtain the likelihood. One difficulty is that the model is not identifiable when $\tau^2=0$ as the t-distribution degrees of freedom doesn't affect the likelihood, and this causes problems with the fitting optimiser. To avoid this a model was fitted with $\nu=0$, to allow $\tau^2=0$, and the likelihood from this model used if it was equal or larger than given by the optimisation with ν unconstrained.

For the robust mixture model a generalized EM (GEM) algorithm was used. Starting values were obtained by partitioning the model into outliers and non-outliers, starting with no outliers. Then each study in turn was selected as an outlier, approximate starting values determined, and the model fitted. The model with the largest likelihood for all models with none or a single outlier was chosen. The process was then repeated adding each possible outlier in turn for two outlier models. This process is repeated until the likelihood cannot be increased, giving the final model. This approach was significantly faster than generating a random selection of outliers. It may be possible to improve the process further by only looking at studies with high residuals as possible outliers.

2 Using package metaplus

The three main procedures available in **metaplus** are `metaplus`, `outlierProbs` and `testOutliers`, with the arguments for each shown in Table 1. The main procedure is

	`metaplus()` arguments
`yi`	Vector of observed effect sizes.
`sei`	Vector of observed standard errors corresponding to each effect size.
`mods`	Data frame of covariates corresponding to each study (only required for a meta-regression model).
`random`	The type of random effect distribution. One of "normal", "t-dist", "mixture", for standard normal, *t*-distribution or mixture of normals respectively.
`label`	The label to be used for this model when producing the summary line on the forest plot. This allows for identification of model when comparing multiple models.
`plotci`	Is the plot obtained from the profiling routines be plotted? See the package **bbmle** for further details.
`justfit`	Should model only be fitted? If not, then profiling and likelihood ratio statistics are not calculated. This is useful for bootstrapping to reduce computation time.
`slab`	Vector of character strings corresponding to each study. This is used to label the plots.

	`outlierProbs()` arguments
`object`	`metaplus` object.

	`testOutliers()` arguments
`object`	`metaplus` object.
`R`	Number of simulations used in the parametric bootstrap.

Table 1: Arguments for procedures of **metaplus** package.

`metaplus` which fits a meta-analysis model to the studies, with results extracted using `summary`, and plotted using `plot`. The `plot` procedure makes use of the `forest` procedure in **metafor** allowing the same customisations of the plots. A parameter specific to `plot` in **metaplus** is `extrameta`, which allows for extra meta-analysis results to be plotted. This allows for different models (i.e. standard and robust) to be compared, or for meta-regression the overall effect at different values of the covariates. The procedure `testOutliers` tests for the presence of outliers for the robust models using the parametric bootstrap. The procedure `outlierProbs` determines the posterior probability of each study being on outlier for the normal mixture model. The returned object has an associated `plot` method to plot the outlier probabilities. The returned results are shown in Table 2.

3 Examples

In the following examples, both the robust options are used to demonstrate the capabilities of the package. In practice a choice will be made to use only one. The advantage of the mixture robust is that it gives the posterior probabilities that each study is an outlier. However, it involves an extra parameter, so may be more unstable when fitting only a small number of studies.

3.1 Intravenous magnesium in acute myocardial infarction

A number of studies have been performed to determine the effectiveness of intravenous magnesium in acute myocardial infarction, and the data obtained from Sterne et al. (2001). The studies have caused considerable controversy, as the results of a single large study ISIS-4 (ISIS-4: Collabarative Group (1995)) contradicts the results of a meta-analysis. Higgins and

	metaplus()
results	Matrix containing columns for estimate, lower and upper 95% confidence interval and p-value. If justfit=TRUE then only the parameter estimates are returned.
yi	Vector of observed effect sizes.
sei	Vector of observed standard errors corresponding to each effect size.
mods	Data frame of covariates corresponding to each study (only returned from a meta-regression model).
fittedmodel	Final model returned from **bbmle**.
justfit	Value of justfit passed to metaplus.
random	Type of random effect.
slab	Vector of character strings corresponding to each study. This is used to label the forrest plot.

	outlierProbs()
outlier.prob	Posterior probability that the study is an outlier.
slab	Label for the study.

	testOutliers()
pvalue	p-value obtained from the parametric bootstrap.
observed	Observed value of the likelihood ratio test statistic.
sims	Simulated values of the test statistic under the null hypothesis.

Table 2: Results reported by procedures of **metaplus** package.

Spiegelhalter (2002) discusses some of the history and some suggested methods from a Bayesian perspective. Woods (2002) comments on the variability between studies due to timing of infusion, and Downing (1999) on the higher level of dose used in ISIS-4, with a more recent meta-analysis by Li et al. (2009). Of interest is whether, given the heterogeneity between studies, the ISIS-4 study is unusual. The data have been obtained in the form of effect estimates and standard errors, but if raw data in the form of number of events per number of patients, then these can be converted using the escalc procedure in the **metafor** package, for example. We can perform the standard random effect meta-analysis, and obtain parameter estimates as follows:

```
> data(mag)
> mag.meta <- metaplus(mag$yi, mag$sei, slab=mag$study)
> summary(mag.meta)

         Est. 95% ci.lb 95% ci.ub   pvalue
muhat -0.7463   -1.2583   -0.3428 0.000501
tau2   0.2540

     logLik      AIC      BIC
  -19.68459 43.36918 44.91436
```

Adding plotci=TRUE will add a plot giving details of the profile confidence intervals, and is shown in Figure 1. When the quadratic approximation to the likelihood holds the shape of the curve will be in the form of a "V". In this case, the shape is not symmetric, so this does not hold. This is confirmed by the lack of symmetry of the confidence interval for muhat.

The forest plot showing the studies and overall effect can be shown using plot(mag.meta), and is shown in Figure 2. The metaplus package makes extensive use of the **metafor** package which allows the parameters for the forest plot in **metafor** to be used when creating forest plots. As the results for the magnesium studies are actually

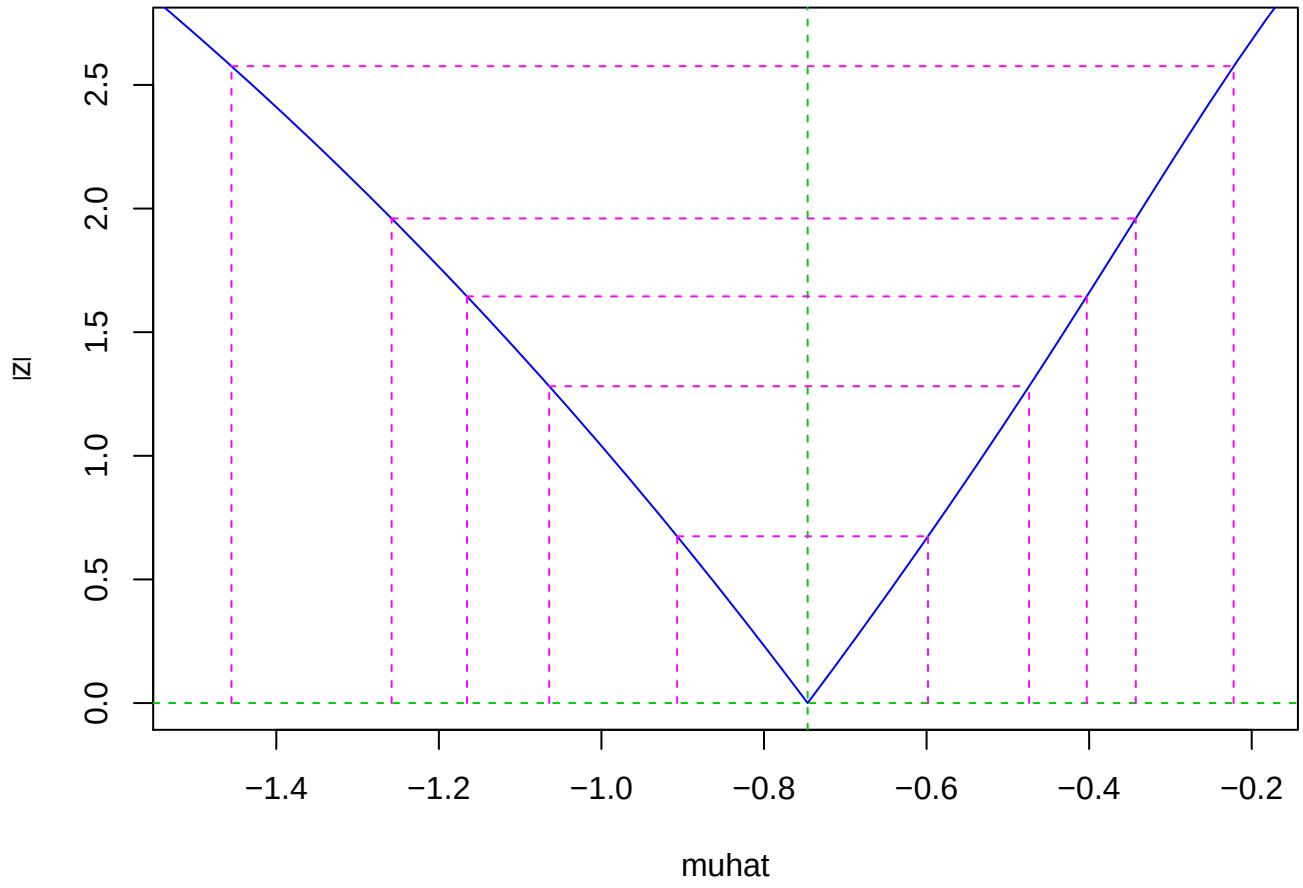

Figure 1: Profile plot.

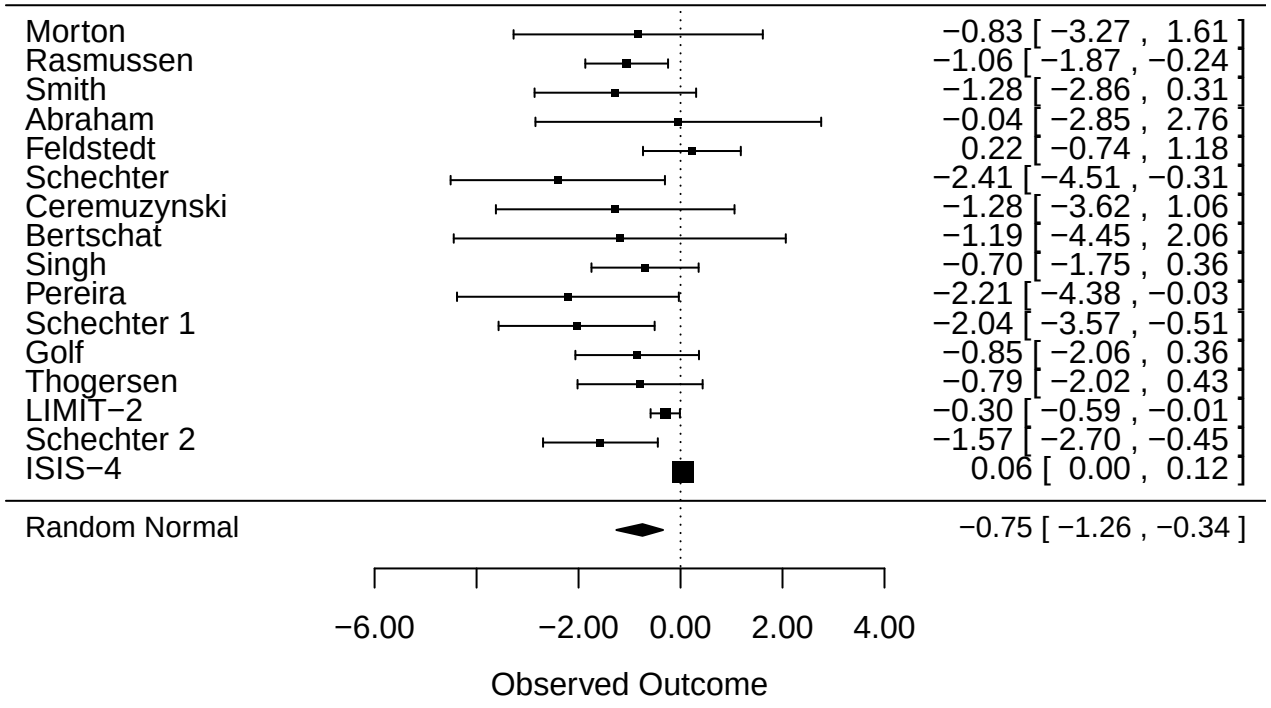

Figure 2: Forest plot for magnesium studies.

log odds ratios it is more useful to produce results with units of odds ratios. This can be obtained by annotating the horizontal axis with odds ratios corresponding to the log odds, and requesting an exponential transformation for the coefficients, as shown in the following code. The plot is shown in Figure 3.

```
> plot(mag.meta, atransf=exp, at=log(c(.01, .1, 1, 10, 100)), xlab="Odds Ratio")
```

Repeating the meta-analysis but now allowing for a *t*-distribution for the random effect by adding the `random="t-dist"` option. From the summary the estimate of `vinv`, the inverse degrees of freedom, is zero indicating infinite degrees of freedom, or a normal distribution. The BIC is also a guide, with an increase for the *t*-distribution model indicating that it is likely that a standard normal is the correct model.

```
> mag.tdist <- metaplus(mag$yi, mag$sei, slab=mag$study,
+         random="t-dist")
> summary(mag.tdist)

         Est. 95% ci.lb 95% ci.ub   pvalue
muhat -0.7463   -1.2583   -0.3430 0.000501
tau2   0.2540
vinv   0.0000

     logLik      AIC      BIC
  -19.68459 45.36918 47.68695
```

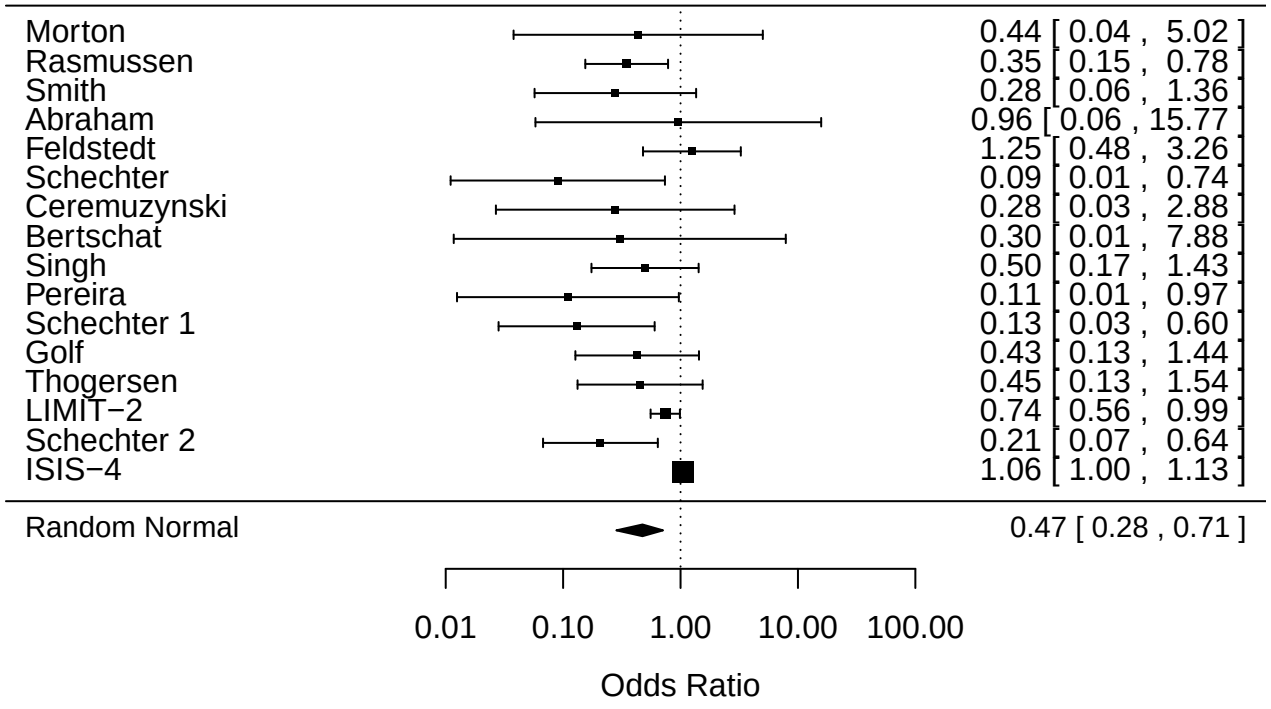

Figure 3: Forest plot for magnesium studies (odds ratios).

We can confirm this with the `testOutliers` command, which performs a parametric bootstrap to obtain the null distribution of the likelihood ratio statistic for the test that $\nu^{-1} = 0$, required as the test is on the boundary of the parameter space. Note that this may take some time for the default of 999 simulations, of the order of one hour or longer depending on the number of studies, so it may be better to start with the parameter `R=99` to limit the number of simulations, with consequently lower accuracy.

```
> summary(testOutliers(mag.tdist))

Observed LRT statistic 0.0 p value 1
```

We can repeat using the robust mixture distribution for the random effect. The variance of both the random effect for standard studies (`tau2`) and for outlier studies (`tau2out`) are very close indicating that there are no outlier studies and this is confirmed by the outlier test.

```
> mag.mix <- metaplus(mag$yi, mag$sei, slab=mag$study,
+       random="mixture")
> summary(mag.mix)

                    Est.  95% ci.lb  95% ci.ub   pvalue
muhat         -0.7463152 -1.2581689 -0.3435347 0.000501
tau2           0.2539984
tau2out        0.2539985
Outlier prob.  0.0006075

     logLik      AIC      BIC
  -19.68459 47.36918 50.45954

> summary(testOutliers(mag.mix))

Observed LRT statistic 0.0 p value 1
```

3.2 CDP choline for cognitive and behavioral disturbances

This meta-analysis evaluates the effect of CDP choline for cognitive and behavioral disturbances associated with chronic cerebral disorders in the elderly (Fioravanti and Yanagi, 2005). A study (Senin 2003) was previously determined to be an outlier by Gumedze and Jackson (2011). We can first fit a standard random effect meta-analysis, as previously.

```
> data(cdp)
> cdp.meta <- metaplus(cdp$yi, cdp$sei, slab=cdp$study)
> summary(cdp.meta)

         Est. 95% ci.lb 95% ci.ub pvalue
muhat 0.38944   0.07269   0.76634 0.0218
tau2  0.14666

     logLik      AIC      BIC
  -8.198544 20.39709 21.00226
```

Note that Senin 2003 does have an unusually high value, as shown in Figure 5. We can then fit a robust model using the t-distribution.

```
> cdp.tdist <- metaplus(cdp$yi, cdp$sei, slab=cdp$study,
+         random="t-dist")
> summary(cdp.tdist)
```

```
          Est. 95% ci.lb 95% ci.ub  pvalue
muhat 1.945e-01 5.294e-02 3.611e-01 0.00899
tau2  4.504e-05
vinv  2.023e+00

     logLik      AIC      BIC
  -4.058334 14.11667 15.02442

> summary(testOutliers(cdp.tdist))

Observed LRT statistic 8.3 p value 0.001
```

As a rough guide, the decrease in AIC and BIC demonstrate that the model is an improvement, and this is confirmed with the outlier test. We can also perform with the robust mixture.

```
> cdp.mix <- metaplus(cdp$yi, cdp$sei, slab=cdp$study,
+         random="mixture")
> summary(cdp.mix)

                Est. 95% ci.lb 95% ci.ub  pvalue
muhat         0.1910    0.0563    0.3479 0.00711
tau2          0.0000
tau2out       3.1558
Outlier prob. 0.1237

     logLik      AIC      BIC
  -3.007145 14.01429 15.22463

> summary(testOutliers(cdp.mix))

Observed LRT statistic 10.4 p value 0.001
```

The output from the robust mixture model has an interesting feature. For standard studies the estimated random effect variance is zero, indicating that only the outlier studies are contributing to the heterogeneity. The posterior probability of each study being an outlier can be obtained as

```
> cdp.mix.outlierProbs <- outlierProbs(cdp.mix)
```

and plotted using `plot(cdp.mix.outlierProbs)` in Figure 4. This shows clearly that Senin 2003 does have a posterior probability of nearly 1.0 of being an outlier. The other studies have a non-zero posterior probability of being outliers, as there is an overlap between the distribution of the standard and outlier studies.

Lastly, we can now produce a forest plot with the results of all three models, using the `extrameta` parameter to add the robust models, as `plot(cdp.meta,extrameta=list(cdp.tdist,cdp.mix))`, and these are shown in Figure 5. The effect of the robust models is to down-weight the Senin 2003 study, which has the consequence of both reducing the overall effect estimate and it's standard error.

3.3 Exercise for depression

This example is a meta-analysis of trials of exercise in the management of depression (Lawlor and Hopker, 2001). Higgins and Thompson (2004) used the data as an example of meta-regression using a number of covariates, which will be limited here to a single covariate, the duration of trial. First perform the meta-analysis using standard normal random effect and the robust mixture model.

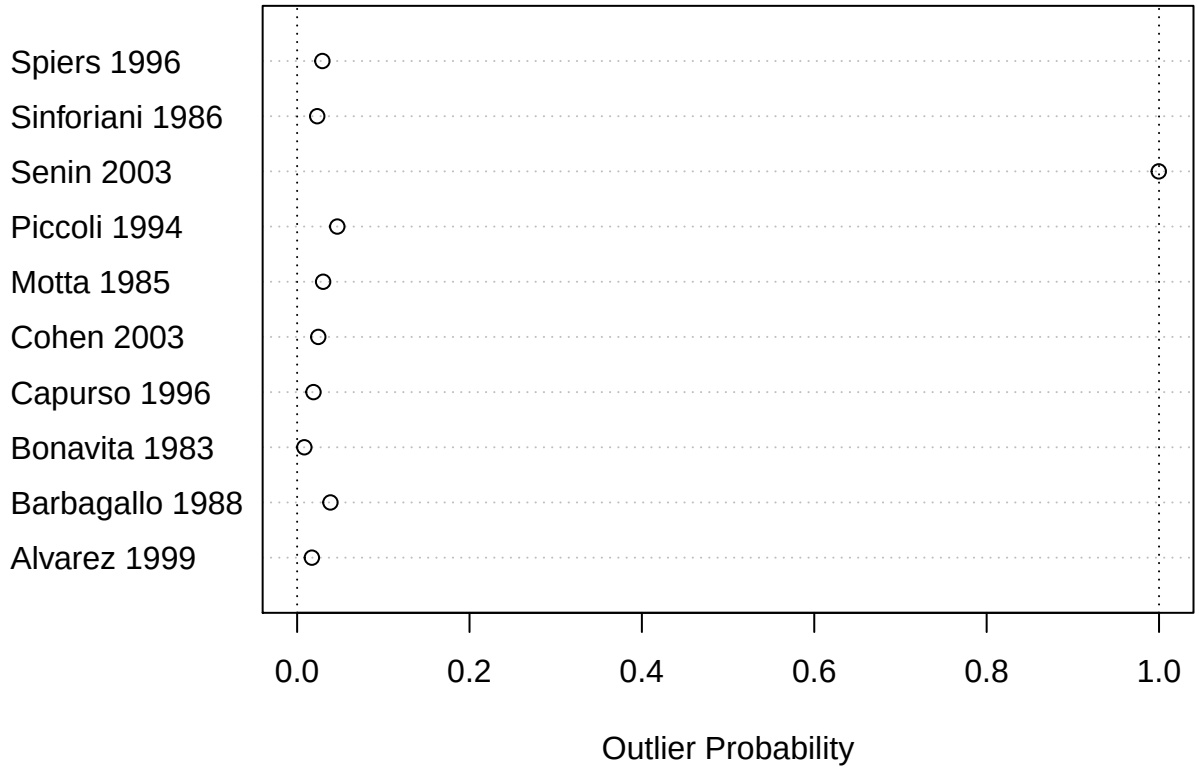

Figure 4: Outlier probabilities for CDP studies.

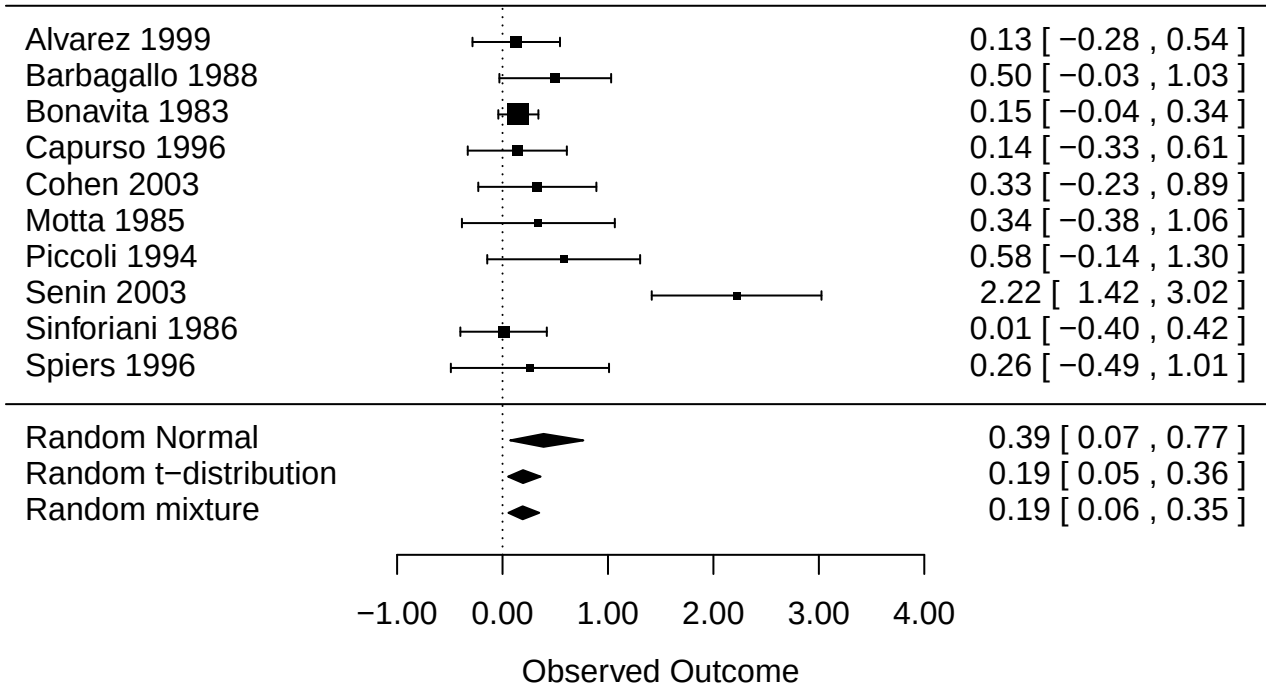

Figure 5: Forest plot for CDP Studies.

```
> data(exercise)
> exercise.meta <- metaplus(exercise$smd, sqrt(exercise$varsmd),
+       mods=exercise[, c("duration"), drop=FALSE], slab=exercise$study)
> summary(exercise.meta)

            Est. 95% ci.lb 95% ci.ub   pvalue
muhat    -2.8994   -4.3006   -1.5222 0.000884
tau2      0.1171
duration  0.2078    0.0584    0.3632 0.011570

     logLik      AIC      BIC
  -8.133435 22.26687 23.17462

> exercise.mix <- metaplus(exercise$smd, sqrt(exercise$varsmd),
+       mods=exercise[, c("duration"), drop=FALSE],
+       slab=exercise$study, random="mixture")
> summary(exercise.mix)

                  Est. 95% ci.lb 95% ci.ub   pvalue
muhat         -2.88472  -4.10690  -1.48262 0.000649
tau2           0.00000
tau2out        0.59398
Outlier prob.  0.25169
duration       0.21086   0.07824   0.34581 0.007123

    logLik      AIC     BIC
  -7.69139 25.38278 26.8957

> exercise.testOutliers <- testOutliers(exercise.mix)
> summary(exercise.testOutliers)

Observed LRT statistic 0.9 p value 0.05

> exercise.outlierProbs <- outlierProbs(exercise.mix)
```

The test for outliers was close to significant (p=0.05), however a conservative approach seems appropriate, by using the robust model where the presence of outliers is not conclusive but there is a reasonable amount of evidence that there are outliers, as in this case. Note also that the p-value is different from that obtained in Beath (2014), due to the use of randomly generated data in the parametric bootstrap. Running the parametric bootstrap with a large number of simulations showed that the p-value was in fact near 0.04. Using `plot(exercise.outlierProbs)` the outlier probabilities are shown in Figure 6 where the study by Reuter is an obvious outlier with a posterior probability greater than 0.9. This study is a dissertation and was not published in a peer-reviewed journal, and was not included in a later meta-analysis by Krogh et al. (2011). There is also strong evidence of the effect of trial duration.

To calculate the effect at each of Weeks 4, 8 and 12 it is easiest to centre the data at those times and fit a meta-regression for each. The intercept for each meta-regression will then be the estimated mean effect at that time. We also fit a model without including the covariate for study duration. The forest plot is shown in Figure 7. This shows that the effect of exercise decreases rapidly the longer the trial, possibly indicating a placebo effect that rapidly wears off. It would also be possible to plot the results from the standard random effect models as well.

```
> exercise$duration4 <- exercise$duration-4
> exercise$duration8 <- exercise$duration-8
```

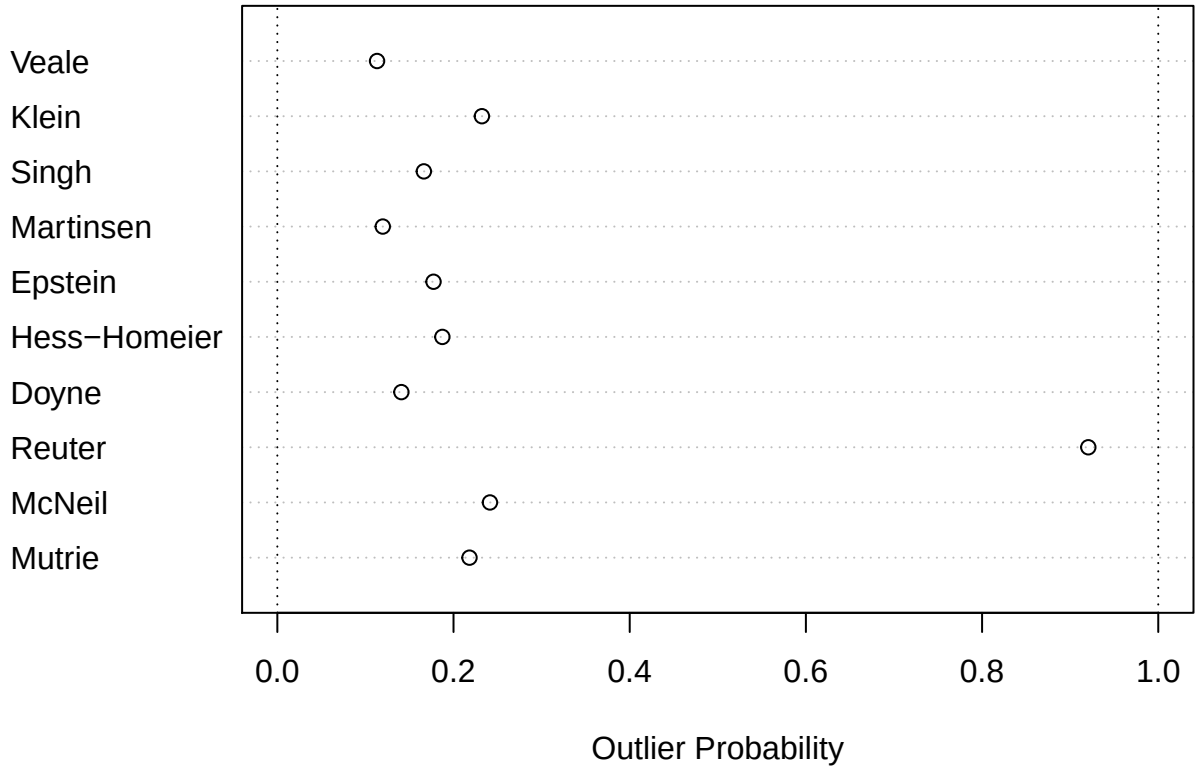

Figure 6: Outcome probabilities for depression versus exercise.

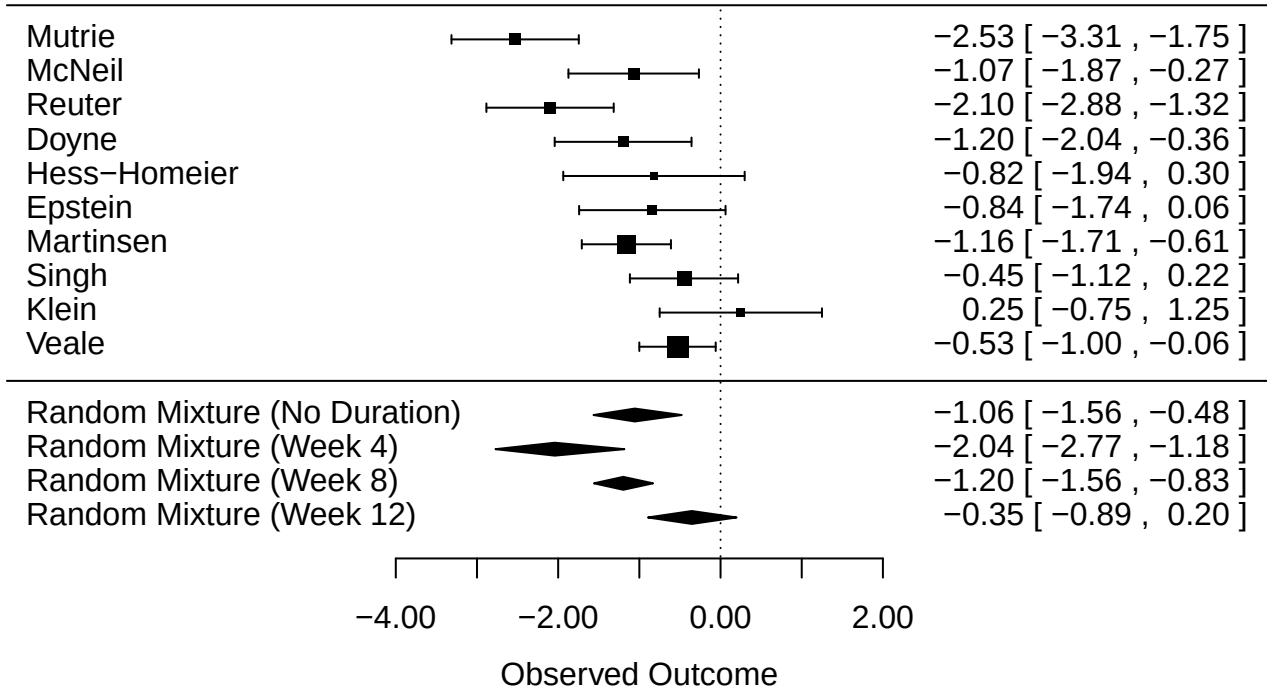

Figure 7: Forest plot for exercise versus depression studies.

```
> exercise$duration12 <- exercise$duration-12
> exercise.nodurn <- metaplus(exercise$smd, sqrt(exercise$varsmd),
+     label="Random Mixture (No Duration)", slab=exercise$study,
+     random="mixture")
> exercise.wk4 <- metaplus(exercise$smd, sqrt(exercise$varsmd),
+     mods=exercise[, c("duration4"), drop=FALSE],
+     label="Random Mixture (Week 4)", slab=exercise$study,
+     random="mixture")
> exercise.wk8 <- metaplus(exercise$smd, sqrt(exercise$varsmd),
+     mods=exercise[, c("duration8"), drop=FALSE],
+     label="Random Mixture (Week 8)", slab=exercise$study,
+     random="mixture")
> exercise.wk12 <- metaplus(exercise$smd, sqrt(exercise$varsmd),
+     mods=exercise[, c("duration12"), drop=FALSE],
+     label="Random Mixture (Week 12)", slab=exercise$study,
+     random="mixture")
> plot(exercise.nodurn, extrameta=list(exercise.wk4, exercise.wk8,
+       exercise.wk12))
```

4 Conclusions and future developments

The capabilities of the **metaplus** package have been presented for fitting both standard normal random effect and robust random effect models. Using a number of examples it has

been shown how it can test for the presence of outliers and compare the results of the robust and standard methods for both meta-analysis and meta-regression.

The design of the package allows for expansion in other areas. A planned future functionality is to fit binary data, using likelihood methods based on distribution of the binomial responses, rather than the log odds ratios fitted using a normal distribution which is the method currently used. The robust methods can then be applied in a similar way to the current models. A possible future expansion is to allow for other robust distributions although this doesn't seem necessary given the simularity of the results obtained in Baker and Jackson (2008) to those using the *t*-distribution.

References

R. Baker and D. Jackson. A new approach to outliers in meta-analysis. *Health Care Management Science*, 11(2):121–131, Dec. 2008. ISSN 1386-9620. [p1, 2, 16]

K. J. Beath. A finite mixture method for outlier detection and robustness in meta-analysis. *Research Synthesis Methods*, 5:285–293, 2014. [p1, 2, 13]

C. Berkey, D. Hoaglin, F. Mosteller, and G. Colditz. A random-effects regression model for meta-analysis. *Statistics in Medicine*, 14:395–411, 1995. [p2]

D. Downing. Is Isis-4 reasearch misconduct? *Journal of Nutritional & Environmental Medicine*, 9:5–13, 1999. [p5]

M. Fioravanti and M. Yanagi. Cytidinediphosphocholine (CDP choline) for cognitive and behavioural disturbances associated with chronic cerebral disorders in the elderly (Review). *The Cochrane Database of Systematic Reviews*, 2005. URL `http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD000269.pub2/pdf`. [p9]

F. N. Gumedze and D. Jackson. A random effects variance shift model for detecting and accommodating outliers in meta-analysis. *BMC Medical Research Methodology*, 11:19, Jan. 2011. ISSN 1471-2288. [p1, 9]

R. J. Hardy and S. G. Thompson. A likelihood approach to meta-analysis with random effects. *Statistics in Medicine*, 15:619–629, 1996. [p3]

J. P. T. Higgins and D. J. Spiegelhalter. Being sceptical about meta-analyses: a Bayesian perspective on magnesium trials in myocardial infarction. *International Journal of Epidemiology*, 31(1):96–104, Feb. 2002. [p4]

J. P. T. Higgins and S. G. Thompson. Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine*, 23(11):1663–82, June 2004. ISSN 0277-6715. [p10]

ISIS-4: Collabarative Group. ISIS-4: A randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58 050 patients with suspected acute myocardial infarction. *The Lancet*, 345(8951):669–682, Mar. 1995. ISSN 01406736. [p4]

J. Krogh, M. Nordentoft, J. A. C. Sterne, and D. A. Lawlor. The effect of exercise in clinically depressed adults: systematic review and meta-analysis of randomized controlled trials. *The Journal of Clinical Psychiatry*, 72(4):529–38, Apr. 2011. ISSN 1555-2101. [p13]

D. A. Lawlor and S. W. Hopker. The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. *British Medical Journal*, 322(31 March):1–8, 2001. [p10]

K. J. Lee and S. G. Thompson. Flexible parametric models for random effects distributions. *Statistics in Medicine*, 27:418–434, 2008. [p1, 2]

J. Li, Q. Zhang, M. Zhang, and M. Egger. Intravenous magnesium for acute myocardial infarction (Review). *The Cochrane Library*, 2009. URL http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD002755.pub2/pdf. [p5]

G. McLachlan. On bootstrapping the likelihood ratio test statistic for the number of components in a normal mixture. *Applied Statistics*, 36(3):318–324, 1987. [p3]

Y. Pawitan. *In All Likelihood: Statistical Modelling and Inference using Likelihood*. Clarendon Press, Oxford, 1st edition, 2001. [p3]

J. A. C. Sterne, M. Bradburn, and M. Egger. Meta-analysis in Stata. In M. Egger, G. Smith, and D. Altman, editors, *Systematic Reviews in Health Care: Meta-Analysis in Context*, chapter 18, pages 347–369. BMJ Publishing Group, 2001. [p4]

W. Viechtbauer. Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, 36(3):1–48, 2010. [p1]

K. L. Woods. Commentary: Biostatistics , biological mechanisms and Bayes: lessons from. *International Journal of Epidemiology*, 31:105–106, 2002. [p5]

Ken J. Beath
Australia
ken@kjbeath.id.au