



## UvA-DARE (Digital Academic Repository)

### Meta-analysis in JASP

Grasman, R.

**Publication date**

2017

**Document Version**

Final published version

[Link to publication](#)

**Citation for published version (APA):**

Grasman, R. (Author). (2017). Meta-analysis in JASP. Web publication/site, JASP. <https://jasp-stats.org/2017/11/15/meta-analysis-jasp/>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



# Meta-analysis in JASP

□ NOVEMBER 15 - 2017

The screenshot shows the JASP software interface. The 'Common' tab is active, and a dropdown menu is open, highlighting 'Meta Analysis'. Below the menu, a table displays meta-analysis data for three studies.

	trial	auth		tpos	
1	1	Aronson	1948	4	119
2	2	Ferguson & Simes	1949	6	300
3	3	Rosenthal et al	1960	3	228

*The JASP meta-analysis module was supported by a SSMART grant from the Berkeley Initiative for Transparency in the Social Sciences (BITSS), an initiative of the Center for Effective Global Action (CEGA).*

The new release of JASP supports an extensive arrange of commonly used techniques for meta-analysis. These include fixed and random effects analysis, fixed and mixed effects meta-regression, forest and funnel plots, tests for funnel plot asymmetry, trim-and-fill and fail-safe N analysis, and more. The engine behind this analysis power is the software developed in the [metafor-project](#). Here we'll give a quick run through of all the functionality currently supported in JASP.

## Example analysis

In the running example below, we'll use the BCG vaccine data set that compiles available evidence for the effectiveness of the BCG vaccine in preventing tuberculosis. The data is available here:

trial	author	year	tpos	tneg	cpos	cneg	ablat	alloc	ES	SE
1	Aronson	1948	4	119	11	128	44	random	-0.9387	0.5976
2	Ferguson & Simes	1949	6	300	29	274	55	random	-1.6662	0.4562
3	Rosenthal et al	1960	3	228	11	209	42	random	-1.3863	0.6583
4	Hart & Sutherland	1977	62	13536	248	12619	52	random	-1.4564	0.1425
5	Frimodt-Moller et al	1973	33	5036	47	5761	13	alternate	-0.2191	0.2279
6	Stein & Aronson	1953	180	1361	372	1079	44	alternate	-0.9581	0.0995
7	Vandiviere et al	1973	8	2537	10	619	19	random	-1.6338	0.4765
8	TPT Madras	1980	505	87886	499	87892	13	random	0.0120	0.0633
9	Coetzee & Berjak	1968	29	7470	45	7232	27	random	-0.4717	0.2387
10	Rosenthal et al	1961	17	1699	65	1600	42	systematic	-1.4012	0.2746
11	Comstock et al	1974	186	50448	141	27197	18	systematic	-0.3408	0.1119
12	Comstock & Webster	1969	5	2493	3	2338	33	systematic	0.4466	0.7309
13	Comstock et al	1976	27	16886	29	17825	33	systematic	-0.0173	0.2676

The data can be downloaded from [here](#).

The columns code meta-information about the publication ( `author` , `year` ), about the reported statistics extracted from the publications ( `tpos` , `tneg` , `cpos` , `cneg` ), and study characteristics ( `ablat` , `alloc` ). The columns `tpos` , `tneg` , `cpos` , `cneg` where the number of tbc tests that came out positive, respectively negative, in the treatment and control condition. The studies varied in patient allocation, which could be either random, alternate, or systematic ( `alloc` ), and varied in geographical latitude of which the absolute value ( `ablat` ) was determined to be evaluated as a moderator of the effect size.

The additional columns `ES` and `SE` are, respectively, the estimated effect size (chosen to be the log odds ratio) for the study, and its standard error. They were computed in a spreadsheet program with the formulas

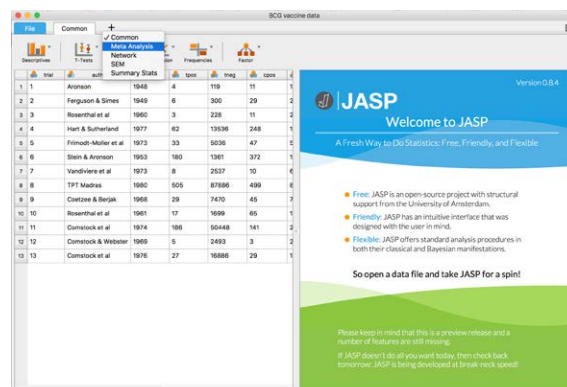
$$ES = \ln \left( \frac{t_{\text{pos}}}{t_{\text{neg}}} \right) - \ln \left( \frac{c_{\text{pos}}}{c_{\text{neg}}} \right)$$

$$SE = \sqrt{\frac{1}{t_{\text{pos}}} + \frac{1}{t_{\text{neg}}} + \frac{1}{c_{\text{pos}}} + \frac{1}{c_{\text{neg}}}}$$

Alternatively, they can be computed with the [effect size tool](#) on the Campbell Collaboration website. A next release of JASP will incorporate a similar effect size calculator tool.

## Opening the meta-analysis module

If we open the data in JASP we can choose the meta-analysis module from the '+' popup menu, and clicking the 'Classical' button. The interface that is displayed is shown below. The main input fields are "Effect Size" and "Effect Size Standard Error". These are required fields; without inserting any of the available variables in these fields the output will be blank as shown. Other important fields are



## Accessing the meta-analysis module

1. "Methods", in which you specify the type of analysis that you want to run. In particular, whether you want to do a Fixed effects analysis or a Random effects analysis, and what type of estimator for the residual heterogeneity variance  $\tau^2$ . The default is a random effects analysis with a restricted maximum likelihood estimator (REML).
2. "Study labels", which allows you to associate a text label with each of the studies.

3. “Covariates” in which you insert any numerical covariates to be included in a meta-regression.
4. “Factors”, in which you insert any nominal scale covariates to be included in a meta-regression.

The screenshot shows the JASP software interface for a meta-analysis of BCG vaccine data. The 'Meta Analysis' tab is selected, and the 'Effect Size' field is set to 'ES' and 'Effect Size Standard Error' to 'SE'. The 'Method' is set to 'Restricted ML'. The 'Results' window displays the following tables:

**Fixed and Random Effects**

	Q	df	p
Omnibus test of Model Coefficients	.	.	.
Test of Residual Heterogeneity	.	.	.

Note. *p*-values are approximate.

**Coefficients**

	Estimate	Standard Error	z	p
	.	.	.	.

Note. Wald test.

**Residual Heterogeneity Estimates**

	Estimate
$\tau^2$	.
$\tau$	.
$I^2$ (%)	.
$H^2$	.

## A simple fixed effects analysis

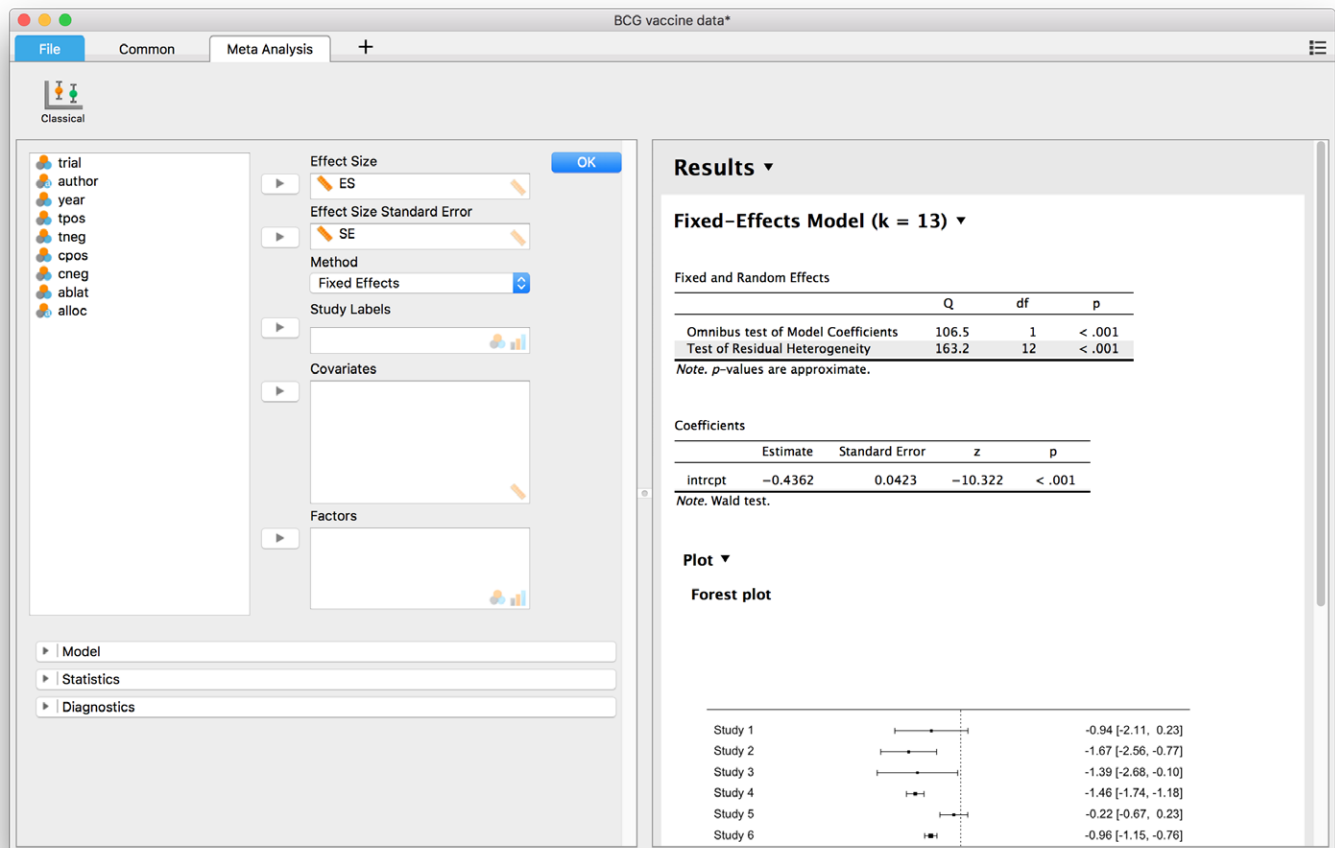
When we drag the **ES** variable into the “Effect Size” field, and the **SE** variable into the “Effect Size Standard Error” field, we instantaneously obtain a result in the output window. In order to ensure that this is a Fixed effects analysis, we have to set the “Methods” dropdown to ‘Fixed effects’. The result is shown below.

In the case of a Fixed effects analysis, the output shows 2 tables. The first table summarises the significance of the effect sizes. The first line labelled ‘Omnibus test of the Model Coefficients’ tests the null hypothesis that all the coefficients in the second table are all zero. The second line tests the null hypothesis that all the effect sizes in the studies are all equal (*homogeneity* of effect sizes). In this case, both are highly significant, with *p*-values below .001, refuting the null hypothesis of no effect, and the hypothesis of homogeneity. The latter implies that the vaccine’s effectiveness differed across studies (*heterogeneity* of the effect sizes). The latter suggests that the Fixed effects model is not an appropriate model to summarise the data, and a better description would result from a random effects model, a model with moderators, or a mixed effects model. However, it’s inference can still be useful, as it indicates that

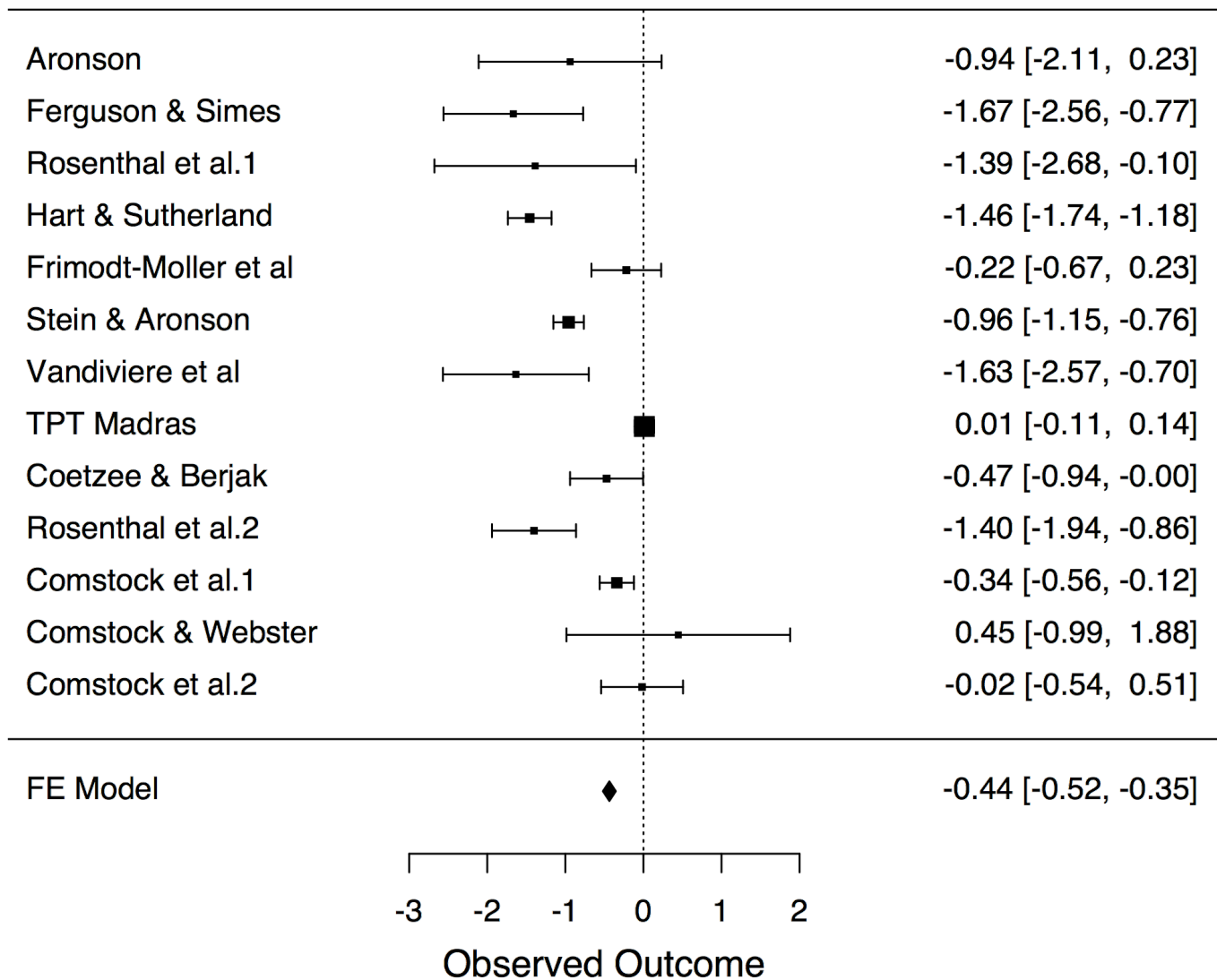
it is safe to conclude that the average of the effect sizes of *the current set of studies* is not equal to zero. (See Hedges & Vevea (1998) for a discussion of fixed vs random effects.)

The second table shows the estimated coefficients for the fixed effects, along with a test of their significance. Here only an intercept is included, the reported coefficient of which is the meta-analytic estimate of the overall (simple) effect size of the combined studies: A weighted average of the effect size estimates in which the weight of each study is the precision of the effect size estimate (the inverse squared standard error). The z-test of the coefficient in this case merely confirms the omnibus test in the first table.

In addition to the tables, by default, a so called forest plot is produced in the output. In a forest plot the effect sizes found for each study are marked with a square mark along the x-axis, one study below the other. The order in which the studies are displayed (currently) follows the order of the data file. The size of the marks is proportional to the weight the study has in determining the combined effect size estimate. In addition to the individual effect size estimates the meta-analytically combined effect size estimate is indicated by the diamond shape at the bottom line of the plot. The width of the diamond indicates the 95% confidence interval for the estimate. The confidence intervals for each study are also indicated in the plot with whiskers. In the case of homogeneity of the effect size, these intervals are expected to cover the combined effect size approximately 95% of the time. The fact that this is not the case here (6 out of these 13 study level confidence intervals do not contain the diamond shape) is yet another indication that the Fixed effects analysis does not give an accurate summary of the effect sizes found in this body of literature. On the right side of the plot these statistics are also displayed numerically. The left hand side shows the study labels, which by default are just the row numbers if no variable was entered in the “Study Labels” field of the interface. The forest plot displayed below was generated by dragging the **author** variable into the “Study Labels” field and saved in PDF format. The forest plot is probably one of the most insightful summary plots of the data in a meta-analysis, and is highly recommended to include in a publication.



*Simple Fixed effects analysis of the BCG vaccine data*



### A simple random effects analysis

Above we concluded that a fixed effects analysis of this set of data does not provide an accurate summary of the current data set. The reason is that the analysis ignores excess variance in the effect sizes that cannot be attributed to sampling error. A (very) simple way to do more justice to the data is to provide an estimate of the excess variance. This variance should be interpreted as systematic unaccounted for differences between studies observed effects. The variance of these differences is often denoted  $\tau^2$ . It turns out that this excess variance also affects the weight that each study should have in the meta-analytically combined effect size estimate. An estimate of  $\tau^2$  and weight adjusted meta-analytic effect



size estimate are obtained in a random effects analysis.

## Setting the Method

In JASP a random effects analysis is obtained by choosing any of the methods other than ‘Fixed Effects’ in the ‘Method’ drop down menu. The most commonly used methods are DerSimonian-Laird, Maximum Likelihood, and Restricted Maximum likelihood. When the focus of the analysis is directed to the fixed effects part of the model, Restricted ML has been recommended. When the estimate of  $\tau^2$  is also of interest, ML has been recommended. The results are displayed below. The key difference with a fixed effects analysis is a third table that gives the estimate of  $\tau^2$ , along with some other metrics that have been found to be useful characterizations of the excess variance. (For instance,  $I^2$  expresses the excess variance as a percentage of the total variance observed in the effect size estimates across studies, and hence is akin an intraclass correlation; for the other measures, please consult the meta-analysis literature.)

BCG vaccine data\*

File Common Meta Analysis +

Classical

trial  
year  
tpos  
tneg  
cpos  
cneg  
ablat  
alloc

Effect Size  
ES

Effect Size Standard Error  
SE

Method  
Restricted ML

Study Labels  
author

Covariates

Factors

OK

Results ▾

**Random-Effects Model (k = 13;  $\tau^2$  estimator: REML) ▾**

Fixed and Random Effects

	Q	df	p
Omnibus test of Model Coefficients	16.05	1	< .001
Test of Residual Heterogeneity	163.19	12	< .001

Note. *p*-values are approximate.

Coefficients

	Estimate	Standard Error	z	p
intrcpt	-0.7452	0.1860	-4.0056	< .001

Note. Wald test.

Residual Heterogeneity Estimates ▾

	Estimate
$\tau^2$	0.3378
$\tau$	0.5812
$I^2$ (%)	92.0747
$H^2$	12.6179

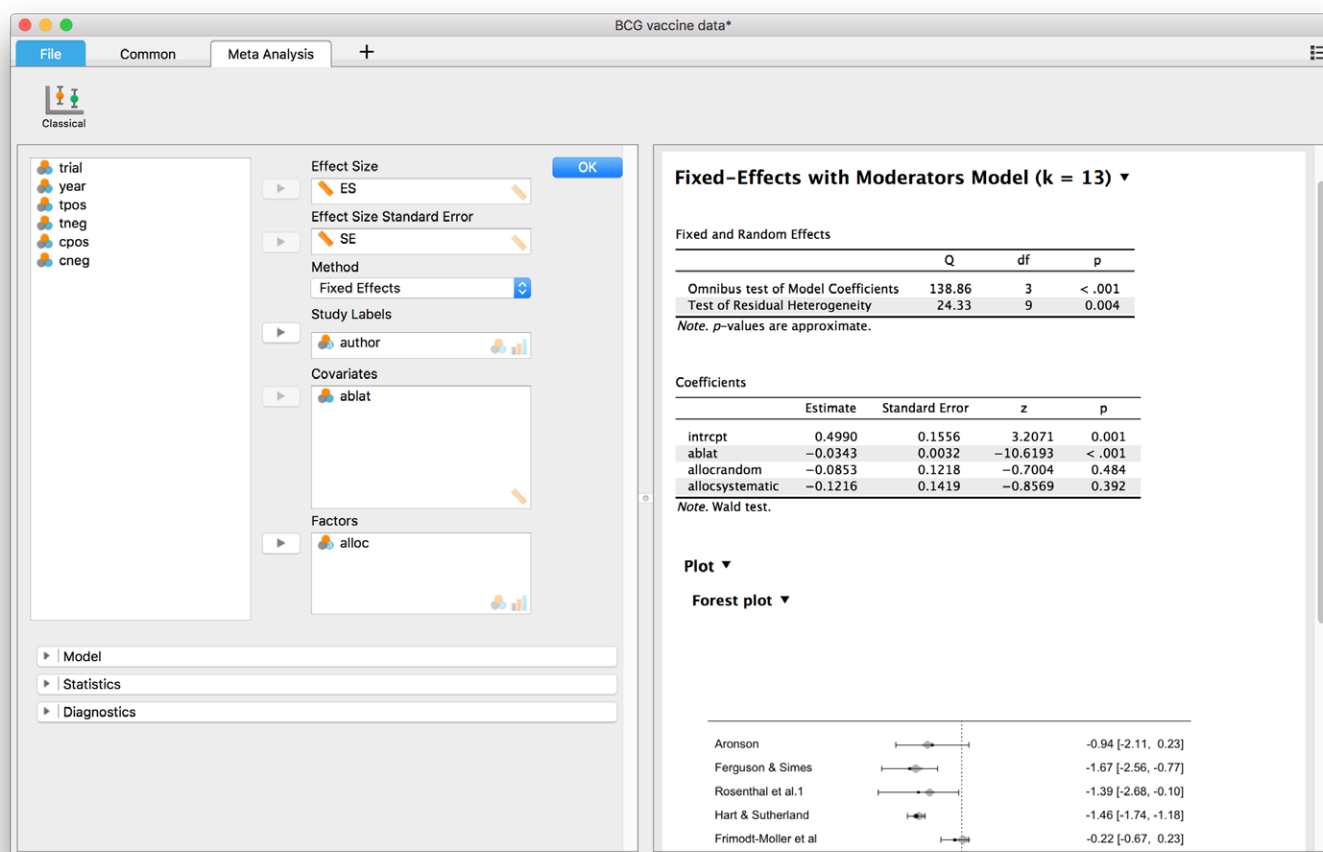
Plot

Forest plot

## A fixed effects meta-regression analysis

In a situation like the current one with the BCG vaccine data set, the random effects model properly makes explicit the excess variance in an estimate of  $\tau^2$ . The random effects model therefore provides a more truthful summary of the effects found in the literature regarding the effectiveness of the vaccine. However, it would be more insightful if we could explain the excess variance. To this end we may try to

explain the differences in the effect between studies, by factors that characterise study specific circumstances and characteristics. Factors that may play a role in the current case are the *geographical latitude* where the vaccine was tested (e.g., because of climatological reasons, economical reasons, or other reasons that are correlated with distance to the equator). Also study design, and in particular the method of *allocation* of study participants among the studies treatment arms, may have significant influence on treatment effects. To see if any of these explain the observed between studies differences, we carry out a meta-regression that includes both the covariate `ablat` as well as the factor `alloc`. We choose the fixed effects method to see if these effect size moderators are sufficient to explain all the excess variance. The results are displayed below.



Next to the intercept in the Coefficients table, we now see one coefficient for `ablat`, and two for `alloc`. The omnibus test indicates that not all 3 of these coefficients are equal to zero. The test for residual heterogeneity in the first table indicates that although the moderators do explain the differences between studies to some extent, there still is unexplained excess variance. The Coefficients table itself indicates that only the coefficient for `ablat` is in fact non-zero. Note that the forest plot has changed as well: It now includes grey diamonds for each study separately. These are the predicted effect sizes for these studies computed from the meta-regression model. Furthermore, the overall meta-analytic effect size has now disappeared: There is no overall meta-analytic effect size anymore, because it varies with

the predictors.

## A Mixed effects analysis

As was the case with the simple fixed effects model, because there still is unexplained heterogeneity, the fixed effects meta-regression is not an accurate summary of the evidence of the vaccine effectiveness conveyed by this set of studies. The unexplained variance must again be made explicit in an estimate of this variance  $\tau^2$ . To this end we change the Method to Restricted Maximum Likelihood which gives an unbiased estimate of  $\tau^2$ . Again this affects the weight that each of the studies carry in the compounded meta-analytic estimated effect.

The screenshot shows the JASP software interface for a meta-analysis of BCG vaccine data. The 'Meta Analysis' tab is active, and the 'Method' is set to 'Restricted ML'. The 'Results' panel displays the following information:

**Mixed-Effects Model (k = 13;  $\tau^2$  estimator: REML)**

Fixed and Random Effects

	Q	df	p
Omnibus test of Model Coefficients	12.50	3	0.006
Test of Residual Heterogeneity	24.33	9	0.004

Note. p-values are approximate.

Coefficients

	Estimate	Standard Error	z	p
intrcpt	0.2610	0.4017	0.6496	0.516
ablat	-0.0294	0.0091	-3.2223	0.001
alloccrandom	-0.1801	0.3482	-0.5171	0.605
alloccsystematic	0.1461	0.3765	0.3881	0.698

Note. Wald test.

Residual Heterogeneity Estimates

	Estimate
$\tau^2$	0.1402
$\tau$	0.3744
$I^2$ (%)	68.7961
$H^2$	3.2047

Plot

Forest plot

## Other options

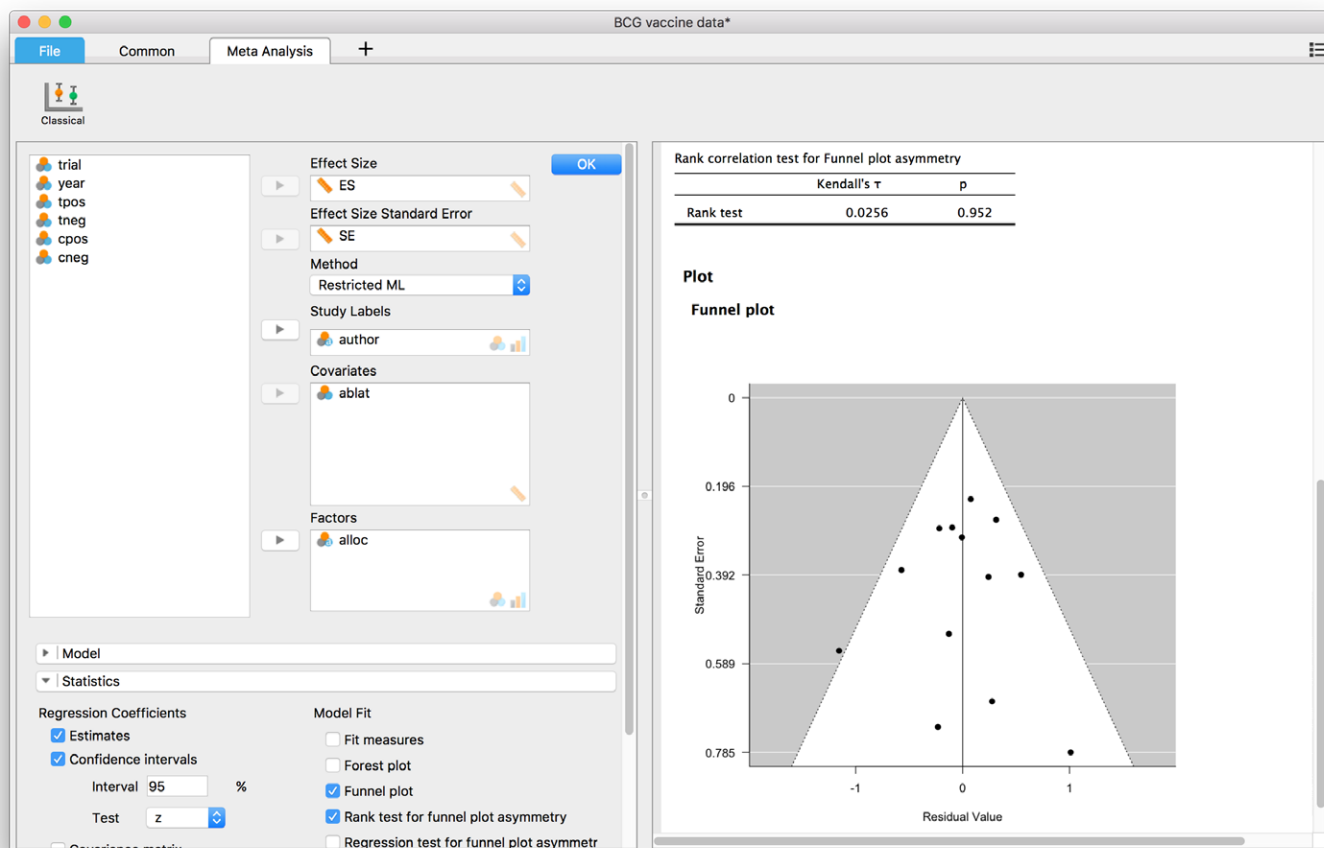
The module allows for a number of other common statistics to be computed and displayed.

- **Model** allows for building models with interactions
- **Statistics** offers a range of statistics and plots that are likely of interest, including statistics commonly used for the assessment of publication bias
- **Diagnostics** offers a range of diagnostics, including robustness analysis tools

## Statistics

Typically confidence intervals are interest, which can be added to the tables by ticking the ‘confidence intervals’ box. Two types are available: Confidence intervals based on z statistic (Gaussian approximation), and confidence intervals based on the adjustment proposed in (Knapp & Hartung, 2003). The former are most conventional.

Another feature of likely interest are the various tools often used for publication bias assessment. The first is a funnel plot, which plots study precision ( $1 / SE$ ) against observed effect size (ES). In general this plot is expected to be symmetrical in the vertical axis around the meta-analytic compounded effect size estimate. Also, the points are expected to lie in a confidence triangle 95% of the time. An excessive amount of points outside the interval can result from heterogeneity in fixed effects models, and non-normal distributed between study differences in random or mixed effects models. Asymmetry of the funnel plot is often interpreted as evidence of publication bias. Funnel plot asymmetry can be formally tested with a non-parametric rank test, or the parametric regression test (also known as “Egger’s test”). In both cases, low p-values are indicative of asymmetry, and by extension, taken as evidence for publication bias.



The running example shows no signs of funnel plot asymmetry. Both visually, as well as according to the rank test.

## Diagnostics

Various diagnostics are available. Besides residuals and influential cases assessment, methods specific to meta-analysis are Fail-safe N and trim-and-fill.

- Fail-safe N estimates the number of zero finding studies needed to add to the list of studies in order to bring the currently observed significance of the meta-analytic effect size to a significant level of  $\alpha = 0.05$ .
- Trim-and-fill tries to add 'missing studies' by adding points that will make the funnel plot symmetric. The added points are then included in the current analysis to adjust the parameter estimates in an attempt to mitigate the influence of publication bias.

Both these methods have been severely criticised and their usefulness is disputed. Below the results for these analysis for the BCG vaccine data are displayed. Fail-safe N is equal to 608, indicating that 608 studies with study effect size 0 would have to be added to the meta-analysis in order to bring the observed significance back to  $\alpha = 0.05$ . The trim-and-fill analysis doesn't change any of the results, because the funnel plot is not found to be asymmetric. The plot produced for trim-and-fill analysis includes several subplots: A forest plot with the adjusted effect size estimates, a funnel plot that would include the estimated 'missing studies' as white (as opposed to black) dots if there were any (not in this example), a radial version of the funnel plot, and a normal q-q plot that can be used to assess departure from normality.

BCG vaccine data\*

File Common Meta Analysis +

Classical

tneg cpos cneg

Effect size Standard Error SE OK

Method Restricted ML

Study Labels author

Covariates ablat

Factors alloc

Model

Statistics

Diagnostics

Model Fit Assessment

- Trim-fill diagnostics
- Casewise diagnostics
- Profile plot
- Fail-safe N
- Diagnostic plots
- Q-Q plot standardized residuals

	Estimate	Lower Bound	Upper Bound
$\tau^2$	0.1402	0.0192	0.9336
$\tau$	0.3744	0.1385	0.9662
$I^2$ (%)	68.7961	23.1741	93.6249
$H^2$	3.2047	1.3016	15.6860

Note. Wald test.

Residual Heterogeneity Estimates

	Estimate	Lower Bound	Upper Bound
$\tau^2$	0.1402	0.0192	0.9336
$\tau$	0.3744	0.1385	0.9662
$I^2$ (%)	68.7961	23.1741	93.6249
$H^2$	3.2047	1.3016	15.6860

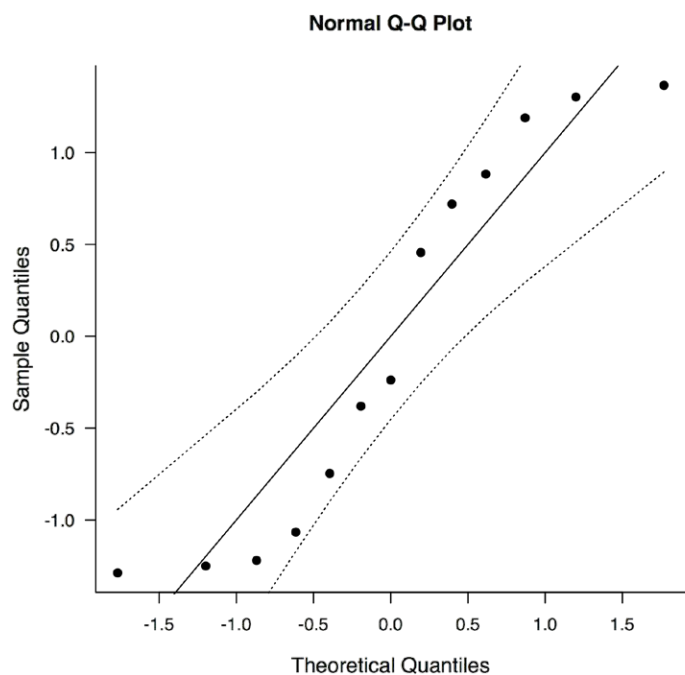
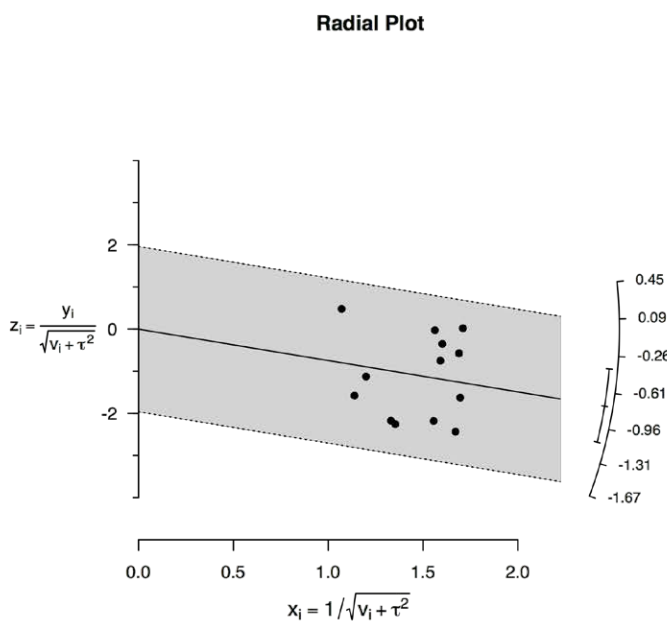
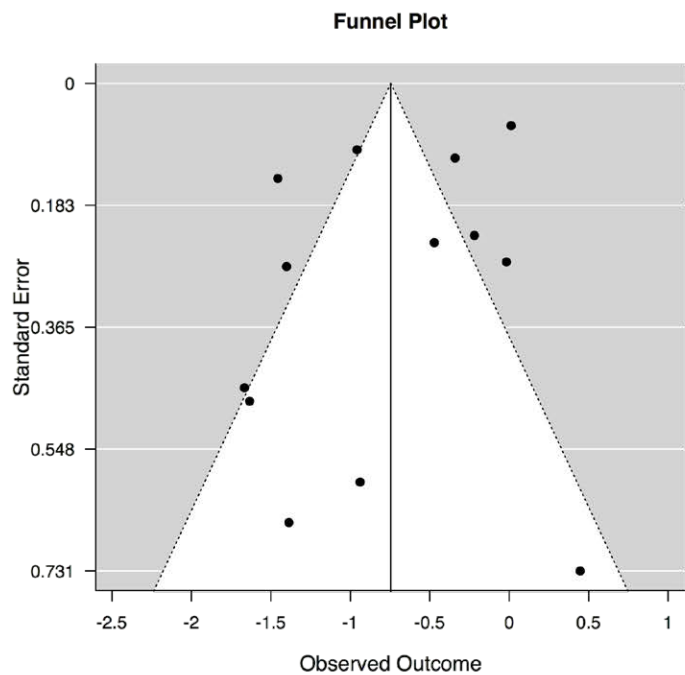
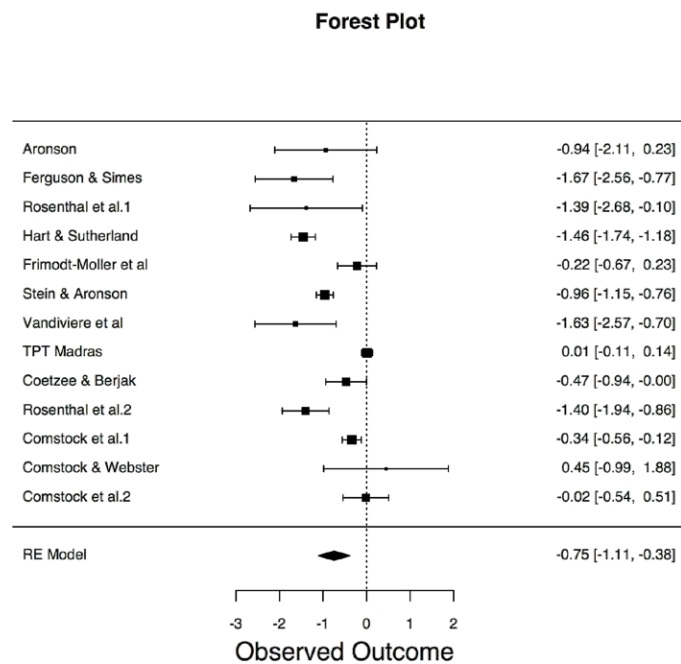
Influence Measures

	Std. Residual	DFFITS	Cook's Distance	Cov. Ratio	$\tau^2(-)$	Q
Aronson	0.4110	0.1640	0.0312	1.6968	0.1612	24
Ferguson & Simes	-0.2296	-0.1128	0.0207	2.3074	0.1694	24
Rosenthal et al.1	-0.3077	-0.0699	0.0063	1.4494	0.1542	24
Hart & Sutherland	0.0028	0.0285	0.0173	4.0291	0.1886	23
Frimodt-Moller et al	-0.2744	-0.2947	0.1216	5.5972	0.1852	22
Stein & Aronson	0.2744	0.3602	0.1834	6.8031	0.1852	22
Vandiviere et al	-2.3839	-0.9889	0.8209	0.1230	0.0649	15
TPT Madras	1.1469	1.1345	1.3452	2.2473	0.1467	16
Coetzee & Berjak	0.5965	0.3844	0.1844	2.8096	0.1822	24
Rosenthal et al.2	-1.5121	-1.0311	1.0305	1.2767	0.1335	22
Comstock et al.1	-0.7072	-0.6423	0.4934	3.1238	0.1683	21
Comstock & Webster	1.3054	0.4078	0.1643	0.8159	0.1273	21
Comstock et al.2*	2.1325	1.5519	1.4656	0.0428	0.0402	15

Note. Cases marked with \* are influential.

File Drawer Analysis

	Fail-safe N	Target Significance	Observed Significance
Rosenthal	608	0.0500	3.3281e -30



Like this post?

Subscribe to our newsletter to receive regular updates about JASP including our latest blog posts, JASP articles, example analyses, new features, interviews with team members, and more! You can unsubscribe at any time.

## ABOUT THE AUTHOR



### Raoul Grasman

Raoul Grasman is associate professor at the Psychological Methods Group at the University of Amsterdam. At JASP, he is responsible for improving code and developing new modules.