Description

`meta summarize` summarizes `meta` data. It reports individual effect sizes and the overall effect size (ES), their confidence intervals (CIs), heterogeneity statistics, and more. `meta summarize` can perform random-effects (RE), common-effect (CE), and fixed-effects (FE) meta-analyses. It can also perform subgroup, cumulative, and sensitivity meta-analyses. For graphical display of meta-analysis summaries, see [META] `meta forestplot`.

Quick start

Perform meta-analysis and summarize `meta` data, which were declared by either `meta set` or `meta esize`

```
meta summarize
```

As above, but summarize meta-analysis results using the empirical Bayes RE method instead of the declared method

```
meta summarize, random(ebayes)
```

As above, but report transformed effect sizes and CIs using the hyperbolic tangent function

```
meta summarize, random(ebayes) transform(tanh)
```

Perform subgroup meta-analysis based on the categorical variable `x1`

```
meta summarize, subgroup(x1)
```

Perform subgroup analysis based on the categorical variables `x1`, `x2`, and `x3`

```
meta summarize, subgroup(x1 x2 x3)
```

Perform cumulative meta-analysis (CMA), where studies are included in the CMA based on the ascending order of observations in variable `x4`

```
meta summarize, cumulative(x4)
```

As above, but stratify the results of the CMA based on groups of the categorical variable `x5`

```
meta summarize, cumulative(x4, by(x5))
```

Perform leave-one-out meta-analysis

```
meta summarize, leaveoneout
```

Perform sensitivity meta-analysis by assuming a fixed value of 0.2 for the between-study heterogeneity parameter $\tau^2$, assuming that the declared model is RE

```
meta summarize, tau2(.2)
```
Menu

Statistics > Meta-analysis

Syntax

Meta-analysis as declared with meta set or meta esize

```
meta summarize [if] [in] [, options reopts ]
```

Random-effects meta-analysis

```
meta summarize [if] [in] , random [remethod] [options reopts ]
```

Common-effect meta-analysis

```
meta summarize [if] [in] , common [cefname] [options]
```

Fixed-effects meta-analysis

```
meta summarize [if] [in] , fixed [cefname] [options]
```

options Description

Main

<table>
<thead>
<tr>
<th>subgroup(varlist)</th>
<th>subgroup meta-analysis for each variable in varlist</th>
</tr>
</thead>
<tbody>
<tr>
<td>cumulative(cumulspec)</td>
<td>cumulative meta-analysis</td>
</tr>
<tr>
<td>leaveoneout</td>
<td>leave-one-out meta-analysis</td>
</tr>
</tbody>
</table>

Options

<table>
<thead>
<tr>
<th>level(#)</th>
<th>set confidence level; default is as declared for meta-analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>eform_option</td>
<td>report exponentiated results</td>
</tr>
<tr>
<td>transform(transfspec)</td>
<td>report transformed results</td>
</tr>
<tr>
<td>sort(varlist[, ...])</td>
<td>sort studies according to varlist</td>
</tr>
<tr>
<td>tdistribution</td>
<td>report t test instead of z test for the overall effect size</td>
</tr>
<tr>
<td>nostudies</td>
<td>suppress output for individual studies</td>
</tr>
<tr>
<td>noheader</td>
<td>suppress output header</td>
</tr>
<tr>
<td>[no] metashow</td>
<td>display or suppress meta settings in the output</td>
</tr>
<tr>
<td>display_options</td>
<td>control column formats</td>
</tr>
</tbody>
</table>

Maximization

<table>
<thead>
<tr>
<th>maximize_options</th>
<th>control the maximization process; seldom used</th>
</tr>
</thead>
</table>

collect is allowed; see [U] 11.1.10 Prefix commands.

remethod Description

<table>
<thead>
<tr>
<th>remdl</th>
<th>restricted maximum likelihood; the default</th>
</tr>
</thead>
<tbody>
<tr>
<td>mle</td>
<td>maximum likelihood</td>
</tr>
<tr>
<td>ebayes</td>
<td>empirical Bayes</td>
</tr>
<tr>
<td>d Laird</td>
<td>DerSimonian–Laird</td>
</tr>
<tr>
<td>sjonkman</td>
<td>Sidik–Jonkman</td>
</tr>
<tr>
<td>hedges</td>
<td>Hedges</td>
</tr>
<tr>
<td>hschmidt</td>
<td>Hunter–Schmidt</td>
</tr>
</tbody>
</table>
### cefemethod

<table>
<thead>
<tr>
<th>cefemethod</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mhaenszel</td>
<td>Mantel–Haenszel</td>
</tr>
<tr>
<td>invvariance</td>
<td>inverse variance</td>
</tr>
<tr>
<td>ivvariance</td>
<td>synonym for invvariance</td>
</tr>
</tbody>
</table>

### reopts

<table>
<thead>
<tr>
<th>reopts</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>tau2(#)</td>
<td>sensitivity meta-analysis using a fixed value of between-study variance $\tau^2$</td>
</tr>
<tr>
<td>i2(#)</td>
<td>sensitivity meta-analysis using a fixed value of heterogeneity statistic $I^2$</td>
</tr>
<tr>
<td>predinterval[(#)]</td>
<td>report prediction interval for the overall effect size</td>
</tr>
<tr>
<td>se(seadj)</td>
<td>adjust standard error of the overall effect size</td>
</tr>
</tbody>
</table>

### Options

Options `random()`, `common()`, and `fixed()`, when specified with `meta summarize`, temporarily override the global model declared by `meta set` or `meta esize` during the computation. Options `random()`, `common()`, and `fixed()` may not be combined. If these options are omitted, the declared meta-analysis model is assumed; see *Declaring a meta-analysis model* in [META meta data]. Also see *Meta-analysis models* in [META Intro].

- `random()` and `random(remethod)` specify that a random-effects model be assumed for meta-analysis; see *Random-effects model* in [META Intro].
  
  `remethod` specifies the type of estimator for the between-study variance $\tau^2$. `remethod` is one of `reml`, `mle`, `ebayes`, `dlaird`, `sjonkman`, `hedges`, or `hschmidt`. `random` is a synonym for `random(reml)`. See *Options* in [META meta esize] for more information.

- `common()` and `common(cefemethod)` specify that a common-effect model be assumed for meta-analysis; see *Common-effect (“fixed-effect”) model* in [META Intro]. Also see the discussion in [META meta data] about common-effect versus fixed-effects models.
  
  `common` implies `common(mhaenszel)` for effect sizes `lnoratio`, `lnrratio`, and `rdiff` and `common(invvariance)` for all other effect sizes. `common(mhaenszel)` is supported only with effect sizes `lnoratio`, `lnrratio`, and `rdiff`.

- `cefemethod` is one of `mhaenszel` or `invvariance` (synonym `ivvariance`). See *Options* in [META meta esize] for more information.

- `fixed()` and `fixed(cefemethod)` specify that a fixed-effects model be assumed for meta-analysis; see *Fixed-effects model* in [META Intro]. Also see the discussion in [META meta data] about fixed-effects versus common-effect models.
  
  `fixed` implies `fixed(mhaenszel)` for effect sizes `lnoratio`, `lnrratio`, and `rdiff` and `fixed(invvariance)` for all other effect sizes. `fixed(mhaenszel)` is supported only with effect sizes `lnoratio`, `lnrratio`, and `rdiff`.

- `cefemethod` is one of `mhaenszel` or `invvariance` (synonym `ivvariance`); see *Options* in [META meta esize] for more information.

- `subgroup(varlist)` specifies that a subgroup meta-analysis (subgroup analysis) be performed for each variable in `varlist`. Subgroup analysis performs meta-analysis separately for each variable in `varlist` and for each group as defined by that variable. The specified meta-analysis model is assumed for
each subgroup. This analysis is useful when the results of all studies are too heterogeneous to be combined into one estimate but the results are similar within certain groups of studies. The specified variables can be numeric or string variables. When multiple variables are specified, only the subgroup results are displayed; that is, the results from individual studies are suppressed for brevity. This option may not be combined with cumulative() or leaveoneout.

cumulative(ordervar[, ascending|descending by(byvar)]) performs a cumulative meta-analysis (CMA). CMA performs multiple meta-analyses and accumulates the results by adding one study at a time to each subsequent analysis. It is useful for monitoring the results of the studies as new studies become available. The studies enter the CMA based on the ordered values of variable ordervar. ordervar must be a numeric variable. By default, ascending order is assumed unless the suboption descending is specified; only one of ascending or descending is allowed. The by(byvar) option specifies that the CMA be stratified by variable byvar. This option may not be combined with subgroup() or leaveoneout.

leaveoneout performs a leave-one-out meta-analysis. For each study, the corresponding leave-one-out meta-analysis is a meta-analysis of all the studies except that study. It is useful for assessing the effect of a single study on the meta-analysis results and for identifying outliers if they exist. This option may not be combined with subgroup() or cumulative().

reopts are tau2(#), i2(#), predinterval[#], and se(khartung[, truncated]). These options are used with random-effects meta-analysis.

tau2(#) specifies the value of the between-study variance parameter, \( \tau^2 \), to use for the random-effects meta-analysis. This option is useful for exploring the sensitivity of the results to different levels of between-study heterogeneity. Only one of tau2() or i2() may be specified. This option is not allowed in combination with subgroup(), cumulative(), or leaveoneout.

i2(#) specifies the value of the heterogeneity statistic \( I^2 \) (as a percentage) to use for the random-effects meta-analysis. This option is useful for exploring the sensitivity of the results to different levels of between-study heterogeneity. Only one of i2() or tau2() may be specified. This option is not allowed in combination with subgroup(), cumulative(), or leaveoneout.

predinterval and predinterval(#) specify that the 95% or #% prediction interval be reported for the overall effect size in addition to the confidence interval. # specifies the confidence level of the prediction interval. The prediction interval provides plausible ranges for the effect size in a future, new study. This option is not allowed in combination with subgroup(), cumulative(), or leaveoneout.

se(seadj) specifies that the adjustment seadj be applied to the standard error of the overall effect size. Additionally, the test of significance of the overall effect size is based on a Student’s \( t \) distribution instead of the normal distribution.

seadj is khartung[, truncated]. Adjustment khartung specifies that the Knapp–Hartung adjustment (Hartung and Knapp 2001a, 2001b; Knapp and Hartung 2003), also known as the Sidik–Jonkman adjustment (Sidik and Jonkman 2002), be applied to the standard error of the overall effect size. hknapp and sjonkman are synonyms for khartung. truncated specifies that the truncated Knapp–Hartung adjustment (Knapp and Hartung 2003), also known as the modified Knapp–Hartung adjustment, be used.

level(#) specifies the confidence level, as a percentage, for confidence intervals. The default is as declared for the meta-analysis session; see Declaring a confidence level for meta-analysis in [META] meta data. Also see option level() in [META] meta set.
eform_option is one of eform, eform(string), or, or rr. It reports exponentiated effect sizes and transforms their respective confidence intervals, whenever applicable. By default, the results are displayed in the metric declared with meta set or meta esize such as log odds-ratios and log risk-ratios. eform_option affects how results are displayed, not how they are estimated and stored.

eform(string) labels the exponentiated effect sizes as string; the other options use default labels. The default label is specific to the chosen effect size. For example, option eform uses Odds ratio when used with log odds-ratios declared with meta esize or Risk ratio when used with the declared log risk-ratios. Option or is a synonym for eform when log odds-ratio is declared, and option rr is a synonym for eform when log risk-ratio is declared. If option eslabel(eslab) is specified during declaration, then eform will use the exp(eslab) label or, if eslab is too long, the exp(ES) label.

transform([ label: ] transf_name) reports transformed effect sizes and CIs. transf_name is one of corr, efficacy, exp, invlogit, or tanh. When label is specified, the transformed effect sizes are labeled as label instead of using the default label. This option may not be combined with eform_option.

corr transforms effect sizes (and CIs) specified as Fisher’s z values into correlations and, by default, labels them as Correlation; that is, transform(corr) is a synonym for transform(Correlation: tanh).

efficacy transforms the effect sizes and CIs using the 1 − exp() function (or more precisely, the −expm1() function) and labels them as Efficacy. This transformation is used, for example, when the effect sizes are log risk-ratios so that the transformed effect sizes can be interpreted as treatment efficacies, 1 − risk ratios.

exp exponentiates effect sizes and CIs and, by default, labels them as exp(ES). This transformation is used, for example, when the effect sizes are log risk-ratios, log odds-ratios, and log hazard-ratios so that the transformed effect sizes can be interpreted as risk ratios, odds ratios, and hazard ratios. If the declared effect sizes are log odds-ratios or log risk-ratios, the default label is Odds ratio or Risk ratio, respectively.

invlogit transforms the effect sizes and CIs using the inverse-logit function, invlogit(), and, by default, labels them as invlogit(ES). This transformation is used, for example, when the effect sizes are logit of proportions so that the transformed effect sizes can be interpreted as proportions.

tanh applies the hyperbolic tangent transformation, tanh(), to the effect sizes and CIs and, by default, labels them as tanh(ES). This transformation is used, for example, when the effect sizes are Fisher’s z values so that the transformed effect sizes can be interpreted as correlations.

sort(varlist[ , ascending | descending ] ) sorts the studies in ascending or descending order based on values of the variables in varlist. This option is useful if you want to sort the studies in the output by effect sizes, sort(_meta_es), or by precision, sort(_meta_se). By default, ascending order is assumed unless the suboption descending is specified; only one of ascending or descending is allowed. varlist may contain string and numeric variables. This option is not allowed with cumulative(). When sort() is not specified, the order of the studies in the output is based on the ascending values of variable _meta_id, which is equivalent to sort(_meta_id).

tdistribution reports a t test instead of a z test for the overall effect size. This option may not be combined with option subgroup(), cumulative(), leaveoneout, or se().
nostudies (synonym nostudy) suppresses the display of information such as effect sizes and their CIs for individual studies from the output table.

noheader suppresses the output header.
meta summarize — Summarize meta-analysis data

metashow and nometashow display or suppress the meta setting information. By default, this information is displayed at the top of the output. You can also specify nometashow with meta update to suppress the meta setting output for the entire meta-analysis session.

display_options: cformat(\%fmt), pformat(\%fmt), and sformat(\%fmt); see [R] Estimation options. The defaults are cformat(\%9.3f), pformat(\%5.3f), and sformat(\%8.2f).

wgtformat(\%fmt) specifies how to format the weight column in the output table. The default is wgtformat(\%5.2f). The maximum format width is 5.

ordformat(\%fmt) specifies the format for the values of the order variable, specified in cumulative(ordervar). The default is ordformat(\%9.0g). The maximum format width is 9.

Maximization

maximize_options: iterate(\#), tolerance(\#), nrtolerance(\#), nonrtolerance (see [R] Maximize), from(\#), and showtrace. These options control the iterative estimation of the between-study variance parameter, \(\tau^2\), with random-effects methods reml, mle, and ebayes. These options are seldom used.

from(\#) specifies the initial value for \(\tau^2\) during estimation. By default, the initial value for \(\tau^2\) is the noniterative Hedges estimator.

showtrace displays the iteration log that contains the estimated parameter \(\tau^2\), its relative difference with the value from the previous iteration, and the scaled gradient.

Remarks and examples

Remarks are presented under the following headings:

Introduction

Examples of using meta summarize

Introduction

Meta-analysis helps answer research questions based on the results of multiple studies. Does exercise prolong life? Does lack of sleep increase the risk of cancer? Does daylight saving save energy? Or does performing the duck-face technique while taking a selfie increase the number of likes on Facebook? These (except perhaps the last one) and many other research questions have been investigated by multiple studies. These studies may have reported conflicting results: some may have shown effects in one direction, some in the opposite, and others may have shown none that are statistically significant. Meta-analysis uses quantitative methods to explore these conflicting results and, whenever possible, provide a unified conclusion based on the results of the individual studies.

Meta-analysis combines the results of similar multiple studies into a single result. Studies typically report some measures of outcomes, or effect sizes, and their precision (standard errors or CIs). Meta-analysis combines the individual effects sizes to provide various meta-analysis summaries. The main summaries are the overall effect size and its precision. Other meta-analysis summaries include the test of significance of the overall effect size, between-study heterogeneity summaries such as the \(I^2\) statistic, and the test of homogeneity between studies. The meta summarize command reports such summaries.
Estimating the overall effect size, $\theta$, and its precision based on the results of multiple studies is at the heart of meta-analysis. There are various methods for estimating $\theta$, which depend on the research goals and model assumptions about the studies. The estimate of the overall (combined) ES is computed as the weighted average of the study-specific effect sizes, with larger weights given to more precise (larger) studies:

$$\hat{\theta} = \frac{\sum_{j=1}^{K} w_j \hat{\theta}_j}{\sum_{j=1}^{K} w_j}$$

The weights are determined by the chosen meta-analysis model, estimation method, and potentially the type of effect size; see Methods and formulas for details. (In [META] Intro, we used $\theta_{pop}$ to denote the population parameter of interest. For simplicity, here and in the rest of the documentation, we will use $\theta$.)

As we described in Meta-analysis models in [META] Intro, the choice of a meta-analysis model is important not only for estimation but also for interpretation of $\hat{\theta}$. meta summarize supports random-effects (random), fixed-effects (fixed), and common-effect (common) meta-analysis models. Each meta-analysis model provides various estimation methods such as the random-effects REML method, random(reml), and fixed-effects Mantel–Haenszel method, fixed(mhaenszel). The default model and method are as declared with meta set or meta esize; see Declaring a meta-analysis model in [META] meta data. Note that the Mantel–Haenszel method is available only with effect sizes lnratio, lnrratio, and rdiff declared by using meta esize; see [META] meta esize.

For random-effects models, you can perform sensitivity meta-analysis to explore the impact of different levels of heterogeneity on the results. You can use the tau2(#) option to specify different fixed values for the between-study variance $\tau^2$. Or you can fix the percentage of variation in the effect sizes because of heterogeneity by specifying the values for the $I^2$ statistic in the i2(#) option. With random-effects models, you can also compute prediction intervals for $\hat{\theta}$, predinterval(#), and use the alternative standard-error estimators, se().

You can perform subgroup analysis, subgroup(), CMA, cumulative(), or leave-one-out meta-analysis, leaveoneout; see Subgroup meta-analysis, Cumulative meta-analysis, and Leave-one-out meta-analysis in [META] Intro. Also see Subgroup meta-analysis, Cumulative meta-analysis, and Leave-one-out meta-analysis in Methods and formulas below.

You can sort the studies based on variables of interest via option sort(). For example, use sort(_meta_es) or sort(_meta_weight) to display the results based on the ascending order of the study effect sizes or study weights, respectively.

You can specify the desired confidence level with level(); report exponentiated results by specifying eform; report a $t$ test, instead of a $z$ test, for the overall effect size by specifying tdistribution; and more.

In the next section, we demonstrate various usages of meta summarize.
Recall the pupil IQ data (Raudenbush and Bryk 1985; Raudenbush 1984) described in *Effects of teacher expectancy on pupil IQ (pupiliq.dta)* of [META] *meta*. Here we will use its declared version and will focus on the demonstration of various options of *meta summarize* and explanation of its output.

```
. use https://www.stata-press.com/data/r17/pupiliqset
(Effects of teacher expectancy on pupil IQ; set with -meta set-)
. keep in 1/10
(9 observations deleted)
. meta query, short
```

For brevity, we consider only the first 10 studies. We use *meta query, short* to remind us about the main settings of the declaration step. Our data were declared by using *meta set* with variables *stdmdiff* and *se* specifying the effect sizes and their standard errors, respectively. The declared meta-analysis model is the default random-effects model with the REML estimation method.

Examples are presented under the following headings:

- Example 1: Default random-effects meta-analysis
- Example 2: DerSimonian–Laird random-effects method
- Example 3: Fixed-effects meta-analysis
- Example 4: Common-effect meta-analysis
- Example 5: Knapp–Hartung standard-error adjustment
- Example 6: Prediction interval
- Example 7: Sensitivity meta-analysis
- Example 8: Other options: CI level, t distribution, sort, eform
- Example 9: Subgroup meta-analysis
- Example 10: Meta-analysis of correlations and the transform() option
- Example 11: Cumulative meta-analysis
- Example 12: Leave-one-out meta-analysis
Example 1: Default random-effects meta-analysis

We type \texttt{meta summarize} to obtain a standard meta-analysis summary.

\begin{verbatim}
.meta summarize

  Effect-size label: Std. mean diff.
  Effect size: stdmdiff
  Std. err.: se
  Study label: studylbl

Meta-analysis summary
Number of studies = 10
Random-effects model
Method: REML
Heterogeneity:
  tau2 = 0.0754
  I2 (%) = 74.98
  H2 = 4.00

Effect size: Std. mean diff.

<table>
<thead>
<tr>
<th>Study</th>
<th>Effect size</th>
<th>[95% conf. interval]</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rosenthal et al., 1974</td>
<td>0.030</td>
<td>-0.215</td>
<td>0.275</td>
</tr>
<tr>
<td>Conn et al., 1968</td>
<td>0.120</td>
<td>-0.168</td>
<td>0.408</td>
</tr>
<tr>
<td>Jose &amp; Cody, 1971</td>
<td>-0.140</td>
<td>-0.467</td>
<td>0.187</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>1.180</td>
<td>0.449</td>
<td>1.911</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>0.260</td>
<td>-0.463</td>
<td>0.983</td>
</tr>
<tr>
<td>Evans &amp; Rosenthal, 1969</td>
<td>-0.060</td>
<td>-0.262</td>
<td>0.142</td>
</tr>
<tr>
<td>Fielder et al., 1971</td>
<td>-0.020</td>
<td>-0.222</td>
<td>0.182</td>
</tr>
<tr>
<td>Claiborn, 1969</td>
<td>-0.320</td>
<td>-0.751</td>
<td>0.111</td>
</tr>
<tr>
<td>Kester, 1969</td>
<td>0.270</td>
<td>-0.051</td>
<td>0.591</td>
</tr>
<tr>
<td>Maxwell, 1970</td>
<td>0.800</td>
<td>0.308</td>
<td>1.292</td>
</tr>
<tr>
<td>theta</td>
<td>0.134</td>
<td>-0.075</td>
<td>0.342</td>
</tr>
</tbody>
</table>

Test of theta = 0: z = 1.26               Prob > |z| = 0.2085
Test of homogeneity: Q = chi2(9) = 26.21  Prob > Q = 0.0019
\end{verbatim}

As with other \texttt{meta} commands, a short information about meta settings is displayed directly following the \texttt{meta summarize} command. It can be suppressed with the \texttt{nometashow} option; see example 2.

Next, the header reports the information about the meta-analysis model and method, the number of studies (10), and several heterogeneity statistics. The output table reports the effect sizes and their 95\% CIs for individual studies and the estimate of the overall, combined ES, labeled as theta, and its 95\% CI. The test of significance of the overall effect size and the homogeneity test are reported at the bottom of the table.

Because our declared effect-size label, \texttt{Std. mean diff.}, was too long to fit as the column header, \texttt{meta summarize} used the generic column label \texttt{Effect size} but displayed the specified label in the table legend.

The mean effect size in our example is 0.134 with the 95\% CI of $[-0.075, 0.342]$. This estimate is computed as the weighted average of the study-specific effect sizes, with the weights representing precision of the studies. The percentages of the total weight for each study are reported in the % weight column. The more precise the study is, the larger its weight percentage. For example, studies 6 and 7, with labels Evans & Rosenthal, 1969 and Fielder et al., 1971, have the largest weight percentage among the studies of about 13\% (each). Thus, their effect-size estimates, $-0.06$ and $-0.02$, have the largest weights in the weighted-average estimate.

The 95\% CI for the overall mean estimate and the test of $H_0: \theta = 0$ with the $z$-test statistic of 1.26 and the $p$-value of 0.2085 suggest that $\theta$ is not statistically significantly different from 0. We should be careful, however, with our conclusions in the presence of between-study heterogeneity.
The heterogeneity statistic $I^2$, reported in the header, is about 75%, which means that 75% of the variability in the effect-size estimates is because of the between-study differences rather than the sampling variation. According to Higgins et al. (2003), this value of $I^2$ corresponds to “large heterogeneity”. (The authors suggest that $I^2 = 25\%$ should indicate “small heterogeneity”, $I^2 = 50\%$ indicate “medium heterogeneity”, and $I^2 = 75\%$ indicate “large heterogeneity”.) The between-study variance $\tau^2$ is estimated to be 0.0754. The homogeneity test of $H_0$: $\theta_1 = \theta_2 = \cdots = \theta_{10}$ reports the $Q$ test statistic of 26.21 with a $p$-value of 0.0019.

When there are few studies, which is typical in meta-analysis, the homogeneity test is known to have low power, which means that it may not detect clinically significant heterogeneity (Hedges and Pigott 2001). Thus, you should use caution when interpreting nonsignificant results as “no heterogeneity”. In fact, many experts (for example, Berman and Parker [2002]) recommend using a 10% significance level instead of the classical 5% level to determine statistical significance when using this test. On the other hand, when there are many studies, this test is known to have excessive power, which means that it tends to detect heterogeneity that is clinically insignificant (Hardy and Thompson 1998).

In our example, the $p$-value of the homogeneity test is 0.0019 < 0.05 < 0.1, so there is definitely statistical evidence of the between-study heterogeneity. See example 9 for one way to account for the heterogeneity.
Example 2: DerSimonian–Laird random-effects method

Continuing with example 1, let's use the DerSimonian–Laird random-effects method instead of the default (declared) REML method. Let’s also suppress the meta setting information displayed at the top of the command output by using the `nometashow` option.

```
.meta summarize, random(dlaird) nometashow
```

Meta-analysis summary
Number of studies = 10
Random-effects model Heterogeneity:
Method: DerSimonian-Laird
\[ \tau^2 = 0.0481 \]
\[ I^2 (\%) = 65.66 \]
\[ H^2 = 2.91 \]

Effect size: Std. mean diff.

<table>
<thead>
<tr>
<th>Study</th>
<th>Effect size</th>
<th>[95% conf. interval]</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rosenthal et al., 1974</td>
<td>0.030</td>
<td>-0.215</td>
<td>0.275</td>
</tr>
<tr>
<td>Conn et al., 1968</td>
<td>0.120</td>
<td>-0.168</td>
<td>0.408</td>
</tr>
<tr>
<td>Jose &amp; Cody, 1971</td>
<td>-0.140</td>
<td>-0.467</td>
<td>0.187</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>1.180</td>
<td>0.449</td>
<td>1.911</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>0.260</td>
<td>-0.463</td>
<td>0.983</td>
</tr>
<tr>
<td>Evans &amp; Rosenthal, 1969</td>
<td>-0.060</td>
<td>-0.262</td>
<td>0.142</td>
</tr>
<tr>
<td>Fielder et al., 1971</td>
<td>-0.020</td>
<td>-0.222</td>
<td>0.182</td>
</tr>
<tr>
<td>Claiborn, 1969</td>
<td>-0.320</td>
<td>-0.751</td>
<td>0.111</td>
</tr>
<tr>
<td>Kester, 1969</td>
<td>0.270</td>
<td>-0.051</td>
<td>0.591</td>
</tr>
<tr>
<td>Maxwell, 1970</td>
<td>0.800</td>
<td>0.308</td>
<td>1.292</td>
</tr>
<tr>
<td>theta</td>
<td>0.117</td>
<td>-0.061</td>
<td>0.296</td>
</tr>
</tbody>
</table>

Test of theta = 0: z = 1.29  Prob > |z| = 0.1967
Test of homogeneity: Q = chi2(9) = 26.21  Prob > Q = 0.0019

The results are now based on the DerSimonian–Laird method, and the header is updated to reflect this. This method is one of the many random-effects methods for estimating the between-study variance \( \tau^2 \). Its estimate is 0.0481. In random-effects models, the weights depend on \( \tau^2 \) and thus will differ across different random-effects methods. The mean effect-size estimate under the DerSimonian–Laird method is 0.117 with the 95% CI of \([-0.061, 0.296]\). This estimate is similar to the 0.134 estimate we obtained in example 1. We also arrive at the same inferential conclusion of no statistical significance of the mean effect size as in the previous example.

To shorten the output, let’s suppress the meta setting information from the output of `meta summarize` for all remaining examples. We can use `meta update` to update our current meta settings.

```
.meta update, nometashow
```

We specified the `nometashow` option with `meta update` to suppress the display of the meta setting information in all `meta` commands; see `Modifying default meta settings` in `[META] meta data`. 

\[ \]
Example 3: Fixed-effects meta-analysis

In example 1, we assumed a random-effects meta-analysis model. We can use the fixed option to specify a fixed-effects meta-analysis model.

```
.meta summarize, fixed
```

<table>
<thead>
<tr>
<th>Study</th>
<th>Effect size</th>
<th>[95% conf. interval]</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rosenthal et al., 1974</td>
<td>0.030</td>
<td>-0.215</td>
<td>0.275</td>
</tr>
<tr>
<td>Conn et al., 1968</td>
<td>0.120</td>
<td>-0.168</td>
<td>0.408</td>
</tr>
<tr>
<td>Jose &amp; Cody, 1971</td>
<td>-0.140</td>
<td>-0.467</td>
<td>0.187</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>1.180</td>
<td>0.449</td>
<td>1.911</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>0.260</td>
<td>-0.463</td>
<td>0.983</td>
</tr>
<tr>
<td>Evans &amp; Rosenthal, 1969</td>
<td>-0.060</td>
<td>-0.262</td>
<td>0.142</td>
</tr>
<tr>
<td>Fielder et al., 1971</td>
<td>-0.020</td>
<td>-0.222</td>
<td>0.182</td>
</tr>
<tr>
<td>Claiborn, 1969</td>
<td>-0.320</td>
<td>-0.751</td>
<td>0.111</td>
</tr>
<tr>
<td>Kester, 1969</td>
<td>0.270</td>
<td>-0.051</td>
<td>0.591</td>
</tr>
<tr>
<td>Maxwell, 1970</td>
<td>0.800</td>
<td>0.308</td>
<td>1.292</td>
</tr>
<tr>
<td>theta</td>
<td>0.051</td>
<td>-0.045</td>
<td>0.146</td>
</tr>
</tbody>
</table>

Test of theta = 0: z = 1.04  Prob > |z| = 0.2974
Test of homogeneity: Q = chi2(9) = 26.21  Prob > Q = 0.0019

As reported in the header, fixed implied the inverse-variance estimation method. The between-group variance parameter is not estimated with fixed-effects models, so the heterogeneity summary does not report tau2. Under this model, the mean effect-size estimate is 0.051 with the 95% CI of \([-0.045, 0.146]\). As we explain in *Comparison between the models and interpretation of their results* in [META] Intro, in a fixed-effects model, theta estimates the weighted average of the true study-specific standardized mean differences. Our interpretation is also limited to these 10 studies that we observed in our meta-analysis. That is, the weighted average of the standardized mean differences of these 10 studies is not statistically significantly different from 0.
Example 4: Common-effect meta-analysis

From example 1 and example 3, we determined that there is substantial between-study variability in these data. Thus, a common-effect model, which assumes that all study-specific effects are the same, is not reasonable for these data. But we will demonstrate it for illustration purposes.

```
.meta summarize, common
```

Meta-analysis summary  Number of studies = 10
Common-effect model
Method: Inverse-variance

<table>
<thead>
<tr>
<th>Effect size: Std. mean diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
</tr>
<tr>
<td>Rosenthal et al., 1974</td>
</tr>
<tr>
<td>Conn et al., 1968</td>
</tr>
<tr>
<td>Jose &amp; Cody, 1971</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
</tr>
<tr>
<td>Evans &amp; Rosenthal, 1969</td>
</tr>
<tr>
<td>Fielder et al., 1971</td>
</tr>
<tr>
<td>Claiborn, 1969</td>
</tr>
<tr>
<td>Kester, 1969</td>
</tr>
<tr>
<td>Maxwell, 1970</td>
</tr>
</tbody>
</table>

theta  0.051 -0.045  0.146

Test of theta = 0: z = 1.04  Prob > |z| = 0.2974

We use the common option to specify a common-effect model. Because this model implies no heterogeneity, the corresponding summaries and the homogeneity test are not reported for this model. As we point out in *Comparison between the models and interpretation of their results* in [META] Intro, a common-effect model is computationally the same as a fixed-effects model. So we obtain the exact same results as in example 3. However, the interpretation of our results is different. Here theta estimates a single effect, which is common to all studies. Although the two models produce the same results, to encourage proper interpretation, we provide both options, common and fixed, to distinguish between these models; see *Declaring a meta-analysis model* in [META] meta data for details.
Example 5: Knapp–Hartung standard-error adjustment

Let's return to our random-effects model from example 1. For random-effects models, meta summarize provides several additional options, which we explore in the next three examples.

The Knapp–Hartung adjustment (also known as the Sidik–Jonkman adjustment) to the standard error of the overall effect size (Knapp and Hartung 2003 and Hartung and Knapp 2001a, 2001b) is sometimes used in practice. We can specify it with the se(khartung) option. We also specify the nostudies option to suppress the output from individual studies because it is unaffected by the se(khartung) option.

```
.meta summarize, se(khartung) nostudies
```

<table>
<thead>
<tr>
<th>Meta-analysis summary</th>
<th>Number of studies = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random-effects model</td>
<td>Heterogeneity:</td>
</tr>
<tr>
<td>Method: REML</td>
<td>tau2 = 0.0754</td>
</tr>
<tr>
<td>SE adjustment: Knapp-Hartung</td>
<td>I2 (%) = 74.98</td>
</tr>
<tr>
<td></td>
<td>H2 = 4.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>theta: Overall Std. mean diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>theta</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>theta</td>
</tr>
</tbody>
</table>

Test of homogeneity: Q = chi2(9) = 26.21 Prob > Q = 0.0019

Without the individual studies, the output table is slightly different. The test of significance is now reported in the output table instead of at the bottom of the output table.

The estimate theta is the same as in example 1, 0.134, but it is reported with more digits in this table. The confidence intervals and the test of significance are different. In addition to making an adjustment to the standard error, Knapp and Hartung also use a Student’s t distribution as a sampling distribution instead of the normal distribution. Thus, the t statistic is reported in the output table instead of the z statistic. Regardless, we still conclude that our overall effect size is not statistically significant.

Another standard error adjustment, also used in practice, is the so-called truncated or modified Knapp–Hartung adjustment; see Methods and formulas for details. This adjustment can be specified with the se(khartung, truncated) option.

```
.meta summarize, se(khartung, truncated)
```

(output omitted)
Example 6: Prediction interval

Recall from Random-effects model in [META] Intro that a random-effects model implies that the observed studies in a meta-analysis represent a sample from a larger population of similar studies. What if we want to estimate the plausible ranges for the overall effect size in a new, future study? We cannot use the confidence interval for the overall effect size because it does not incorporate the uncertainty in estimating the between-study variance, which is important if we want to predict an effect in a new study. We can compute the prediction interval.

```
. meta summarize, predinterval(90) nostudies
```

```
<table>
<thead>
<tr>
<th>Method: REML</th>
<th>tau2 = 0.0754</th>
</tr>
</thead>
<tbody>
<tr>
<td>I2 (%) = 74.98</td>
<td></td>
</tr>
<tr>
<td>H2 = 4.00</td>
<td></td>
</tr>
</tbody>
</table>
```

```
theta: Overall Std. mean diff.  
| Estimate | Std. err. | z  | P>|z| | [95% conf. interval] |
|----------|-----------|----|------|-----------------------|
| 0.1335309 | 0.1061617 | 1.26 | 0.208 | -0.0745422 - 0.3416041 |
```

90% prediction interval for theta: [-0.414, 0.681]

Test of homogeneity: Q = chi2(9) = 26.21 Prob > Q = 0.0019

We specified predinterval(90) to compute the 90% prediction interval for the mean effect size; use predinterval to compute the 95% interval. Following example 5, we also used nostudies to suppress individual studies.

The 90% prediction interval, reported at the bottom of the table, is [-0.414, 0.681]. The prediction interval will be wider than the confidence interval because it additionally accounts for the uncertainty in the between-study variability.

Example 7: Sensitivity meta-analysis

For random-effects models, we can perform sensitivity analysis to explore various levels of heterogeneity between studies. Let’s see how our results change for different values of the between-study variance $\tau^2$ and the heterogeneity statistic $I^2$.

Let’s compute the results assuming that $\tau^2$ equals 0.25.

```
. meta summarize, tau2(0.25) nostudies
```

```
<table>
<thead>
<tr>
<th>Method: User-specified tau2</th>
<th>tau2 = 0.2500</th>
</tr>
</thead>
<tbody>
<tr>
<td>I2 (%) = 90.86</td>
<td></td>
</tr>
<tr>
<td>H2 = 10.94</td>
<td></td>
</tr>
</tbody>
</table>
```

```
theta: Overall Std. mean diff.  
| Estimate | Std. err. | z  | P>|z| | [95% conf. interval] |
|----------|-----------|----|------|-----------------------|
| 0.173588 | 0.171407 | 1.01 | 0.311 | -1.623636 - 0.5095395 |
```

Test of homogeneity: Q = chi2(9) = 26.21 Prob > Q = 0.0019

Our estimate of the mean effect size is 0.174 with the 95% CI of [-0.162, 0.51] compared with 0.134 with the 95% CI of [-0.075, 0.342] from example 1.
The specified value of \( \tau^2 \) corresponds to an \( I^2 \) of about 91%. Let's now compute the results assuming \( I^2 \) of 10%.

```
. meta summarize, i2(10) nostudies
```

Sensitivity meta-analysis summary

Random-effects model

Method: User-specified I2

| theta: Overall Std. mean diff. |
|------------------|--------|---------|--------|--------|--------|--------|
| Estimate | Std. err. | z | P>|z| | [95% conf. interval] |
| theta | .0589369 | .0527232 | 1.12 | 0.264 | -0.0443987 | .1622724 |

Test of homogeneity: Q = chi2(9) = 26.21 Prob > Q = 0.0019

The estimate of the mean effect size is 0.059 with the 95% CI of \([-0.044, 0.162]\). The corresponding \( \tau^2 \) value is 0.0028.

In both cases above, the mean effect size is not statistically significant.

Example 8: Other options: CI level, t distribution, sort, eform

`meta summarize` provides other options such as `level()` to temporarily change the declared confidence level and `tdistribution` to use a Student’s \( t \) distribution as the sampling distribution instead of the default normal distribution.

```
. meta summarize, level(90) tdistribution
```

Meta-analysis summary

Random-effects model

Method: REML

| Effect size: Std. mean diff. |
|------------------|--------|---------|--------|--------|--------|
| Study | Effect size | [90% conf. interval] | % weight |
| Rosenthal et al., 1974 | 0.030 | -0.176 | 0.236 | 12.39 |
| Conn et al., 1968 | 0.120 | -0.122 | 0.362 | 11.62 |
| Jose & Cody, 1971 | -0.140 | -0.415 | 0.135 | 10.92 |
| Pellegrini & Hicks, 1972 | 1.180 | 0.566 | 1.794 | 5.25 |
| Pellegrini & Hicks, 1972 | 0.260 | -0.347 | 0.867 | 5.33 |
| Evans & Rosenthal, 1969 | -0.060 | -0.229 | 0.109 | 13.11 |
| Fielder et al., 1971 | -0.020 | -0.189 | 0.149 | 13.11 |
| Claiborn, 1969 | -0.320 | -0.682 | 0.042 | 9.11 |
| Kester, 1969 | 0.270 | 0.000 | 0.540 | 11.02 |
| Maxwell, 1970 | 0.800 | 0.387 | 1.213 | 8.15 |
| theta | 0.134 | -0.061 | 0.328 |

Test of theta = 0: t(9) = 1.26 Prob > |t| = 0.2085
Test of homogeneity: Q = chi2(9) = 26.21 Prob > Q = 0.0019

Notice that all CIs, including those for the individual studies, now correspond to the 90% confidence level, compared with example 1. Also, the significance test now uses the Student’s \( t \) distribution with 9 degrees of freedom, but the conclusion remains the same—the mean effect size is not statistically significant.
You may also find `meta summarize`’s option `eform` useful when dealing with the effect sizes in the log-transformed metric such as log odds-ratios or log risk-ratios. By default, `meta summarize` reports results in the declared metric, which should be chosen such that the sampling distributions of the effect sizes are well approximated by normal distributions. It may be more convenient, however, to display the final results in the original metric. When you specify the `eform` option, it reports the exponentiated results and the corresponding CIs. Note that the significance tests and other summary measures are still computed based on the nonexponentiated results.

It does not make sense to exponentiate standardized mean differences in our example, but we will do this just to demonstrate the option.

We will also use the `sort()` option to sort our results based on the descending order of study weights, with larger, more precise studies appearing first.

```
.meta summarize, eform sort(_meta_weight, descending)
```

Meta-analysis summary

Number of studies = 10

Random-effects model

Heterogeneity:

Method: REML

tau2 = 0.0754

I2 (%) = 74.98

H2 = 4.00

<table>
<thead>
<tr>
<th>Study</th>
<th>exp(ES)</th>
<th>[95% conf. interval]</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evans &amp; Rosenthal, 1969</td>
<td>0.942</td>
<td>(0.770, 1.152)</td>
<td>13.11</td>
</tr>
<tr>
<td>Fielder et al., 1971</td>
<td>0.980</td>
<td>(0.801, 1.199)</td>
<td>13.11</td>
</tr>
<tr>
<td>Rosenthal et al., 1974</td>
<td>1.030</td>
<td>(0.807, 1.317)</td>
<td>12.39</td>
</tr>
<tr>
<td>Conn et al., 1968</td>
<td>1.127</td>
<td>(0.845, 1.504)</td>
<td>11.62</td>
</tr>
<tr>
<td>Kester, 1969</td>
<td>1.310</td>
<td>(0.950, 1.807)</td>
<td>11.02</td>
</tr>
<tr>
<td>Jose &amp; Cody, 1971</td>
<td>0.869</td>
<td>(0.627, 1.206)</td>
<td>10.92</td>
</tr>
<tr>
<td>Conn, 1969</td>
<td>0.726</td>
<td>(0.472, 1.118)</td>
<td>9.11</td>
</tr>
<tr>
<td>Maxwell, 1970</td>
<td>2.226</td>
<td>(1.361, 3.640)</td>
<td>8.15</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>1.297</td>
<td>(0.629, 2.673)</td>
<td>5.33</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>3.254</td>
<td>(1.567, 6.760)</td>
<td>5.25</td>
</tr>
</tbody>
</table>

Sorted by: _meta_weight

Test of theta = 0: z = 1.26  Prob > |z| = 0.2085
Test of homogeneity: Q = chi2(9) = 26.21  Prob > Q = 0.0019

`meta summarize, eform` reports exponentiated effect sizes and their corresponding CIs. It labels the effect-size column as `exp(ES)`, but you can change this label to `string` by specifying `eform(string)`.

Note that the `eform` option worked in our example because `meta set` declared our precomputed effect sizes as generic. They could have been log odds-ratios, in which case `eform` would make perfect sense. However, if you use `meta esize` to compute the standardized mean differences (for example, Hedges’s g) and try to use `eform` with `meta summarize`, you will receive an error message because `meta summarize` knows that exponentiation is not appropriate with effect sizes that correspond to continuous data. With effect sizes `lnoratio` (or `lnorpeto`) and `lnrratio` computed by `meta esize`, you can also use the respective options `or` and `rr`, which are synonyms for `eform` in those cases. These options (and `eform`) will label your results as Odds ratio (Peto’s OR) and Risk ratio.
### Example 9: Subgroup meta-analysis

In example 1 and example 3, we identified the presence of substantial heterogeneity between the observed studies. Sometimes, the heterogeneity can be explained by some study-level covariates, also known as moderators. With categorical moderators, we can perform subgroup analysis, which performs meta-analysis separately for each category of each moderator.

We have binary variable `week1`, which records whether teachers had prior contact with students for more than 1 week or for 1 week or less. Let's use this variable as the moderator in our subgroup analysis. We specify the variable `week1` in the `subgroup()` option.

```stata
.meta summarize, subgroup(week1)
```

<table>
<thead>
<tr>
<th>Subgroup meta-analysis summary</th>
<th>Number of studies = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random-effects model</td>
<td></td>
</tr>
<tr>
<td>Method: REML</td>
<td></td>
</tr>
<tr>
<td>Group: week1</td>
<td></td>
</tr>
</tbody>
</table>

**Effect size: Std. mean diff.**

<table>
<thead>
<tr>
<th>Study</th>
<th>Effect size</th>
<th>[95% conf. interval]</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group: &lt;= 1 week</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>1.180</td>
<td>0.449 - 1.911</td>
<td>5.25</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>0.260</td>
<td>-0.463 - 0.983</td>
<td>5.33</td>
</tr>
<tr>
<td>Kester, 1969</td>
<td>0.270</td>
<td>-0.051 - 0.591</td>
<td>11.02</td>
</tr>
<tr>
<td>Maxwell, 1970</td>
<td>0.800</td>
<td>0.308 - 1.292</td>
<td>8.15</td>
</tr>
<tr>
<td>theta</td>
<td>0.581</td>
<td>0.174 - 0.989</td>
<td></td>
</tr>
<tr>
<td>Group: &gt; 1 week</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosenthal et al., 1974</td>
<td>0.030</td>
<td>-0.215 - 0.275</td>
<td>12.39</td>
</tr>
<tr>
<td>Conn et al., 1968</td>
<td>0.120</td>
<td>-0.168 - 0.408</td>
<td>11.62</td>
</tr>
<tr>
<td>Jose &amp; Cody, 1971</td>
<td>-0.140</td>
<td>-0.467 - 0.187</td>
<td>10.92</td>
</tr>
<tr>
<td>Evans &amp; Rosenthal, 1969</td>
<td>-0.060</td>
<td>-0.262 - 0.142</td>
<td>13.11</td>
</tr>
<tr>
<td>Fielder et al., 1971</td>
<td>-0.020</td>
<td>-0.222 - 0.182</td>
<td>13.11</td>
</tr>
<tr>
<td>Claiborn, 1969</td>
<td>-0.320</td>
<td>-0.751 - 0.111</td>
<td>9.11</td>
</tr>
<tr>
<td>theta</td>
<td>-0.033</td>
<td>-0.137 - 0.071</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>theta</td>
<td>0.134</td>
<td>-0.075 - 0.342</td>
<td>0.342</td>
</tr>
</tbody>
</table>

**Heterogeneity summary**

<table>
<thead>
<tr>
<th>Group</th>
<th>df</th>
<th>Q</th>
<th>P &gt; Q</th>
<th>tau2</th>
<th>% I²</th>
<th>H²</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;= 1 week</td>
<td>3</td>
<td>7.14</td>
<td>0.068</td>
<td>0.095</td>
<td>57.03</td>
<td>2.33</td>
</tr>
<tr>
<td>&gt; 1 week</td>
<td>5</td>
<td>3.53</td>
<td>0.618</td>
<td>0.000</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Overall</td>
<td>9</td>
<td>26.21</td>
<td>0.002</td>
<td>0.075</td>
<td>74.98</td>
<td>4.00</td>
</tr>
</tbody>
</table>

Test of group differences: $Q_b = \chi^2(1) = 8.18$  Prob > $Q_b = 0.004$

We now have two output tables. Our main table now reports results from individual studies separately for each group, in addition to the group-specific overall effect size. The overall effect size computed using all studies is reported at the bottom under **Overall**.

The second table reports the group-specific and overall heterogeneity summaries. The test of group differences is reported at the bottom of this table.

The estimated theta for the group with contact <= 1 week is 0.581 with the 95% CI of [0.174, 0.989]. The mean effect size in this group is statistically significant at the 5% level. The estimated theta...
for the group with contact > 1 week is $-0.033$ with the 95% CI of $[-0.137, 0.071]$. The mean effect size in this group is not statistically significant at the 5% level.

If we look at the heterogeneity summaries, the $\leq 1$ week group still has some unexplained between-study heterogeneity with an estimated $I^2$ of 57% and a $p$-value of the homogeneity test of $0.068 < 0.1$. There does not appear to be any between-study heterogeneity in the > 1 week group: $I^2$ is essentially 0%, and the homogeneity test $p$-value is 0.618.

We should interpret our results with caution because each subgroup analysis used a few studies, with the $\leq 1$ week group having only 4 studies.

We can specify multiple variables in the `subgroup()` option. Let’s also include variable `tester` in our subgroup analysis.

```
.meta summarize, subgroup(week1 tester)
```

<table>
<thead>
<tr>
<th>Group</th>
<th>No. of studies</th>
<th>Std. mean diff.</th>
<th>[95% conf. interval]</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>week1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;= 1 week</td>
<td>4</td>
<td>0.581</td>
<td>0.174</td>
<td>0.989</td>
</tr>
<tr>
<td>&gt; 1 week</td>
<td>6</td>
<td>-0.033</td>
<td>-0.137</td>
<td>0.071</td>
</tr>
<tr>
<td>tester</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aware</td>
<td>7</td>
<td>0.059</td>
<td>-0.129</td>
<td>0.247</td>
</tr>
<tr>
<td>blind</td>
<td>3</td>
<td>0.316</td>
<td>-0.206</td>
<td>0.837</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>theta</td>
<td>10</td>
<td>0.134</td>
<td>-0.075</td>
<td>0.342</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>df</th>
<th>Q</th>
<th>P &gt; Q</th>
<th>tau2</th>
<th>% I2</th>
<th>H2</th>
</tr>
</thead>
<tbody>
<tr>
<td>week1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;= 1 week</td>
<td>3</td>
<td>7.14</td>
<td>0.068</td>
<td>0.095</td>
<td>57.03</td>
<td>2.33</td>
</tr>
<tr>
<td>&gt; 1 week</td>
<td>5</td>
<td>3.53</td>
<td>0.618</td>
<td>0.000</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>tester</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aware</td>
<td>6</td>
<td>16.35</td>
<td>0.012</td>
<td>0.035</td>
<td>59.07</td>
<td>2.44</td>
</tr>
<tr>
<td>blind</td>
<td>2</td>
<td>9.31</td>
<td>0.009</td>
<td>0.154</td>
<td>75.14</td>
<td>4.02</td>
</tr>
<tr>
<td>Overall</td>
<td>9</td>
<td>26.21</td>
<td>0.002</td>
<td>0.075</td>
<td>74.98</td>
<td>4.00</td>
</tr>
</tbody>
</table>

With more than one variable in `subgroup()`, `meta summarize` reports three output tables. To conserve space, the main table does not report individual studies but reports the number of studies in each group. It also reports the $p$-values of the corresponding significance tests of the overall effect sizes in each group.
The heterogeneity table reports the group summaries for each variable, in addition to the overall summaries. The new table reports the results of tests of subgroup differences for each variable.

The studies appear to be homogeneous across the levels of the tester variable.

Example 10: Meta-analysis of correlations and the transform() option

Molloy, O’Carroll, and Ferguson (2013) conducted a meta-analysis to examine to what degree conscientiousness is related to medication adherence. Medication adherence is the extent to which typically chronically ill patients follow medical recommendations as prescribed. Conscientiousness is defined as “socially prescribed impulse control that facilitates task- and goal-directed behavior, such as thinking before acting, delaying gratification, following norms and rules, and planning, organizing and prioritizing tasks” (John and Srivastava 1999, 121).

The dataset contains the variables studylbl, rho, and n to indicate the authors and year of publication of the studies, the correlation coefficient between conscientiousness and medication adherence, and the study sample size, respectively.

```
. use https://www.stata-press.com/data/r17/adherence
(Conscientiousness and medication adherence)
. describe n rho studylbl
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>Storage</th>
<th>Display</th>
<th>Value label</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>int</td>
<td>%9.0g</td>
<td>Sample size of the study</td>
</tr>
<tr>
<td>rho</td>
<td>double</td>
<td>%9.0g</td>
<td>* Correlation coefficient</td>
</tr>
<tr>
<td>studylbl</td>
<td>str26</td>
<td>%26s</td>
<td>Study label</td>
</tr>
</tbody>
</table>

The correlation coefficient rho is measured on the natural scale ($-1 \leq r \leq 1$), so the first step is to transform rho using the Fisher’s $z$ transformation as follows:

$$fisherz = \frac{1}{2} \log \left( \frac{1 + rho}{1 - rho} \right) = \text{atanh}(rho) \sim N \left( 0, \frac{1}{n - 3} \right)$$

If the underlying data are bivariate normal, the variance of fisherz equals $1/(n-3)$ and depends only on the within-study sample size and not on the correlation parameter itself. Fisher’s $z$ transformation is available in Stata using the atanh() function. Below, we also generate the se variable to contain the values of the (asymptotic) standard errors of the Fisher’s $z$ values in each study and use meta set to declare our meta data.
. generate double fisherz = atanh(rho)
. generate double se = sqrt(1/(n-3))
. meta set fisherz se, studylabel(studylbl) nometashow

Meta-analysis setting information

Study information
  No. of studies: 16
  Study label: studylbl
  Study size: N/A

Effect size
  Type: <generic>
  Label: Effect size
  Variable: fisherz

Precision
  Std. err.: se
  CI: [_meta_cil, _meta_ciu]
  CI level: 95%

Model and method
  Model: Random effects
  Method: REML

The meta-analysis summary may be obtained as follows:

. meta summarize

Meta-analysis summary

Random-effects model
  Number of studies = 16
  Heterogeneity:
  Method: REML
    tau2 = 0.0081
    I2 (%) = 61.73
    H2 = 2.61

<table>
<thead>
<tr>
<th>Study</th>
<th>Effect size</th>
<th>[95% conf. interval]</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axelsson et al. (2009)</td>
<td>0.189</td>
<td>-0.001</td>
<td>0.380</td>
</tr>
<tr>
<td>Axelsson et al. (2011)</td>
<td>0.163</td>
<td>0.092</td>
<td>0.235</td>
</tr>
<tr>
<td>Bruce et al. (2010)</td>
<td>0.354</td>
<td>0.082</td>
<td>0.626</td>
</tr>
<tr>
<td>Christensen et al. (1999)</td>
<td>0.332</td>
<td>0.139</td>
<td>0.524</td>
</tr>
<tr>
<td>Christensen &amp; Smith (1995)</td>
<td>0.277</td>
<td>0.041</td>
<td>0.513</td>
</tr>
<tr>
<td>Cohen et al. (2004)</td>
<td>0.000</td>
<td>-0.249</td>
<td>0.249</td>
</tr>
<tr>
<td>Dobbels et al. (2005)</td>
<td>0.177</td>
<td>0.027</td>
<td>0.327</td>
</tr>
<tr>
<td>Ediger et al. (2007)</td>
<td>0.050</td>
<td>-0.059</td>
<td>0.159</td>
</tr>
<tr>
<td>Insel et al. (2006)</td>
<td>0.266</td>
<td>0.002</td>
<td>0.530</td>
</tr>
<tr>
<td>Jerant et al. (2011)</td>
<td>0.010</td>
<td>-0.061</td>
<td>0.081</td>
</tr>
<tr>
<td>Moran et al. (1997)</td>
<td>-0.090</td>
<td>-0.359</td>
<td>0.179</td>
</tr>
<tr>
<td>O’Cleirigh et al. (2007)</td>
<td>0.388</td>
<td>0.179</td>
<td>0.597</td>
</tr>
<tr>
<td>Penedo et al. (2003)</td>
<td>0.000</td>
<td>-0.184</td>
<td>0.184</td>
</tr>
<tr>
<td>Quine et al. (2012)</td>
<td>0.151</td>
<td>0.066</td>
<td>0.236</td>
</tr>
<tr>
<td>Stilley et al. (2004)</td>
<td>0.245</td>
<td>0.087</td>
<td>0.402</td>
</tr>
<tr>
<td>Wiebe &amp; Christensen (1997)</td>
<td>0.040</td>
<td>-0.209</td>
<td>0.289</td>
</tr>
</tbody>
</table>

| theta                  | 0.150       | 0.088                | 0.212    |

Test of theta = 0: z = 4.75    Prob > |z| = 0.0000
Test of homogeneity: Q = chi2(15) = 38.16    Prob > Q = 0.0009

The overall Fisher’s z value (transformed correlation coefficient) across the 16 studies is estimated to be 0.150 using the REML RE meta-analysis model.

The interpretation of the results, however, is easier in the natural correlation-coefficient metric, which we can compute using the inverse transformation:

\[
rho = \frac{\exp(2 \times \text{fisherz}) - 1}{\exp(2 \times \text{fisherz}) + 1} = \tanh(\text{fisherz})
\]
Thus, you may obtain the value of the correlation coefficient and its CI by typing

```stata
. display tanh(r(theta))
.14880413
. display "[" tanh(r(ci_lb)) ", " tanh(r(ci_ub)) "]"
[.08783366, .20866384]
```

More conveniently, you can use the `transform(corr)` option to report correlations. This option applies the hyperbolic tangent (\( \tanh() \)) transformation to the Fisher’s \( z \) values and labels the resulting effect sizes as Correlation. Notice that specifying `transform(corr)` is equivalent to specifying `transform(Correlation: tanh)`.

```stata
. meta summarize, transform(corr)
```

<table>
<thead>
<tr>
<th>Study</th>
<th>Correlation</th>
<th>[95% conf. interval]</th>
<th>% weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axelsson et al. (2009)</td>
<td>0.187</td>
<td>-0.001</td>
<td>0.362</td>
</tr>
<tr>
<td>Axelsson et al. (2011)</td>
<td>0.162</td>
<td>0.091</td>
<td>0.231</td>
</tr>
<tr>
<td>Bruce et al. (2010)</td>
<td>0.340</td>
<td>0.082</td>
<td>0.555</td>
</tr>
<tr>
<td>Christensen et al. (1999)</td>
<td>0.320</td>
<td>0.139</td>
<td>0.481</td>
</tr>
<tr>
<td>Christensen &amp; Smith (1995)</td>
<td>0.270</td>
<td>0.041</td>
<td>0.472</td>
</tr>
<tr>
<td>Cohen et al. (2004)</td>
<td>0.000</td>
<td>-0.244</td>
<td>0.244</td>
</tr>
<tr>
<td>Dobbels et al. (2005)</td>
<td>0.175</td>
<td>0.027</td>
<td>0.316</td>
</tr>
<tr>
<td>Ediger et al. (2007)</td>
<td>0.050</td>
<td>-0.059</td>
<td>0.158</td>
</tr>
<tr>
<td>Insel et al. (2006)</td>
<td>0.260</td>
<td>0.002</td>
<td>0.486</td>
</tr>
<tr>
<td>Jerant et al. (2011)</td>
<td>0.010</td>
<td>-0.061</td>
<td>0.081</td>
</tr>
<tr>
<td>Moran et al. (1997)</td>
<td>-0.090</td>
<td>-0.345</td>
<td>0.177</td>
</tr>
<tr>
<td>O’Cleirigh et al. (2007)</td>
<td>0.370</td>
<td>0.178</td>
<td>0.535</td>
</tr>
<tr>
<td>Penedo et al. (2003)</td>
<td>0.000</td>
<td>-0.182</td>
<td>0.182</td>
</tr>
<tr>
<td>Quine et al. (2012)</td>
<td>0.150</td>
<td>0.066</td>
<td>0.232</td>
</tr>
<tr>
<td>Stilley et al. (2004)</td>
<td>0.240</td>
<td>0.087</td>
<td>0.382</td>
</tr>
<tr>
<td>Wiebe &amp; Christensen (1997)</td>
<td>0.040</td>
<td>-0.206</td>
<td>0.281</td>
</tr>
</tbody>
</table>

\( \tanh(\text{theta}) \) 0.149 0.088 0.209

Test of \( \theta = 0 \): \( z = 4.75 \) \( \text{Prob} > |z| = 0.0000 \)
Test of homogeneity: \( Q = \chi^2(15) = 38.16 \) \( \text{Prob} > Q = 0.0009 \)

The overall correlation value is 0.149 with a CI of [0.088, 0.209].

### Example 11: Cumulative meta-analysis

CMA (Lau et al. 1992; Sterne 2016) performs multiple meta-analyses by accumulating studies one at a time. The studies are first ordered with respect to a variable of interest, the ordering variable. Meta-analysis summaries are then computed for the first study, for the first two studies, for the first three studies, and so on. The last meta-analysis will correspond to the standard meta-analysis using all studies.

CMA is useful, for instance, for identifying the point in time of the potential change in the direction or significance of the effect size when the ordering variable is time. You can use the `cumulative()` option to perform CMA.
For demonstration purposes, let’s continue with the dataset in example 1 and use year as our ordering variable.

```
.meta summarize, cumulative(year)
```

Cumulative meta-analysis summary
Number of studies = 10
Random-effects model
Method: REML
Order variable: year

<table>
<thead>
<tr>
<th>Study</th>
<th>theta</th>
<th>[95% conf. interval]</th>
<th>p-value</th>
<th>year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conn et al., 1968</td>
<td>0.120</td>
<td>-0.168</td>
<td>0.408</td>
<td>0.414</td>
</tr>
<tr>
<td>Evans &amp; Rosent, 1969</td>
<td>-0.001</td>
<td>-0.166</td>
<td>0.165</td>
<td>0.995</td>
</tr>
<tr>
<td>Claiborn, 1969</td>
<td>-0.042</td>
<td>-0.201</td>
<td>0.117</td>
<td>0.605</td>
</tr>
<tr>
<td>Kester, 1969</td>
<td>0.022</td>
<td>-0.177</td>
<td>0.221</td>
<td>0.830</td>
</tr>
<tr>
<td>Maxwell, 1970</td>
<td>0.140</td>
<td>-0.178</td>
<td>0.459</td>
<td>0.389</td>
</tr>
<tr>
<td>Jose &amp; Cody, 1971</td>
<td>0.089</td>
<td>-0.177</td>
<td>0.355</td>
<td>0.510</td>
</tr>
<tr>
<td>Fielder et al., 1-1</td>
<td>0.064</td>
<td>-0.141</td>
<td>0.270</td>
<td>0.539</td>
</tr>
<tr>
<td>Pellegrini &amp; H, 1972</td>
<td>0.161</td>
<td>-0.117</td>
<td>0.438</td>
<td>0.257</td>
</tr>
<tr>
<td>Pellegrini &amp; H, 1972</td>
<td>0.161</td>
<td>-0.090</td>
<td>0.413</td>
<td>0.208</td>
</tr>
<tr>
<td>Rosenthal et al., 1-4</td>
<td>0.134</td>
<td>-0.075</td>
<td>0.342</td>
<td>0.208</td>
</tr>
</tbody>
</table>

The output table reports the overall effect size and its CIs for each cumulative analysis. The p-value column contains the p-values of the significance tests of the overall effect sizes from these analyses. The last column displays the values of the ordering variable.

In our example, no particular trend is apparent.

We can perform stratified CMA by specifying a categorical variable in cumulative()’s option by(). To demonstrate, we also specify cumulative()’s option descending to list results in descending order of year.

```
.meta summarize, cumulative(year, by(week1) descending)
```

Stratified cumulative meta-analysis summary
Number of studies = 10
Random-effects model
Method: REML
Order variable: year (descending)
Stratum: week1

<table>
<thead>
<tr>
<th>Study</th>
<th>theta</th>
<th>[95% conf. interval]</th>
<th>p-value</th>
<th>year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group: &lt;= 1 week</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pellegrini &amp; H, 1972</td>
<td>0.260</td>
<td>-0.463</td>
<td>0.983</td>
<td>0.481</td>
</tr>
<tr>
<td>Pellegrini &amp; H, 1972</td>
<td>0.718</td>
<td>-0.183</td>
<td>1.620</td>
<td>0.118</td>
</tr>
<tr>
<td>Maxwell, 1970</td>
<td>0.755</td>
<td>0.320</td>
<td>1.190</td>
<td>0.001</td>
</tr>
<tr>
<td>Kester, 1969</td>
<td>0.581</td>
<td>0.174</td>
<td>0.989</td>
<td>0.005</td>
</tr>
<tr>
<td>Group: &gt; 1 week</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosenthal et al., 1-4</td>
<td>0.030</td>
<td>-0.215</td>
<td>0.275</td>
<td>0.810</td>
</tr>
<tr>
<td>Fielder et al., 1-1</td>
<td>0.000</td>
<td>-0.156</td>
<td>0.156</td>
<td>0.998</td>
</tr>
<tr>
<td>Jose &amp; Cody, 1971</td>
<td>-0.026</td>
<td>-0.166</td>
<td>0.115</td>
<td>0.720</td>
</tr>
<tr>
<td>Claiborn, 1969</td>
<td>-0.054</td>
<td>-0.188</td>
<td>0.080</td>
<td>0.429</td>
</tr>
<tr>
<td>Evans &amp; Rosent, 1969</td>
<td>-0.056</td>
<td>-0.167</td>
<td>0.056</td>
<td>0.326</td>
</tr>
<tr>
<td>Conn et al., 1968</td>
<td>-0.033</td>
<td>-0.137</td>
<td>0.071</td>
<td>0.535</td>
</tr>
</tbody>
</table>

CMA is performed separately for each group of week1.
Also see *Cumulative meta-analysis* in [META] *meta*.

### Example 12: Leave-one-out meta-analysis

For each study in the meta-analysis, the corresponding leave-one-out meta-analysis will omit that study and perform a meta-analysis on the remaining set of studies ($k - 1$ studies). It is useful for exploring the influence of a single study on the overall effect size estimate.

Continuing with example 1, we will use option `leaveoneout` to perform a leave-one-out meta-analysis and sort our results according to variable `se` so that larger studies appear first.

```
.meta summarize, leaveoneout sort(se)
```

**Leave-one-out meta-analysis summary**

<table>
<thead>
<tr>
<th>Omitted study</th>
<th>theta</th>
<th>[95% conf. interval]</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evans &amp; Rosenthal, 1969</td>
<td>0.172</td>
<td>-0.073</td>
<td>0.418</td>
</tr>
<tr>
<td>Fielder et al., 1971</td>
<td>0.168</td>
<td>-0.081</td>
<td>0.418</td>
</tr>
<tr>
<td>Rosenthal et al., 1974</td>
<td>0.161</td>
<td>-0.090</td>
<td>0.413</td>
</tr>
<tr>
<td>Conn et al., 1968</td>
<td>0.149</td>
<td>-0.102</td>
<td>0.400</td>
</tr>
<tr>
<td>Kester, 1969</td>
<td>0.127</td>
<td>-0.115</td>
<td>0.368</td>
</tr>
<tr>
<td>Jose &amp; Cody, 1971</td>
<td>0.174</td>
<td>-0.060</td>
<td>0.408</td>
</tr>
<tr>
<td>Claiborn, 1969</td>
<td>0.175</td>
<td>-0.036</td>
<td>0.386</td>
</tr>
<tr>
<td>Maxwell, 1970</td>
<td>0.021</td>
<td>-0.076</td>
<td>0.119</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>0.132</td>
<td>-0.095</td>
<td>0.358</td>
</tr>
<tr>
<td>Pellegrini &amp; Hicks, 1972</td>
<td>0.057</td>
<td>-0.090</td>
<td>0.204</td>
</tr>
<tr>
<td>theta</td>
<td>0.134</td>
<td>-0.075</td>
<td>0.342</td>
</tr>
</tbody>
</table>

Sorted by: `se`

The output table reports the overall effect size and its CIs for each leave-one-out analysis. In this example, the first row reports the overall effect size estimate based on all the studies excluding the Evans & Rosenthal, 1969 study ($10 - 1 = 9$ studies). The p-value column contains the p-values of the significance tests of the overall effect sizes from these analyses. The last row displays the results based on all 10 studies. It seems that the Maxwell, 1970 study has a relatively large influence because the 95% CI from the meta-analysis excluding that study, $[-0.076, 0.119]$, does not contain the overall effect size estimate based on all studies, 0.134.
**Stored results**

`meta summarize` stores the following in `r()`:

**Scalars**
- `r(N)` number of observations
- `r(theta)` overall effect size
- `r(se)` standard error of overall effect size
- `r(ci_lb)` lower CI bound for overall effect size
- `r(ci_ub)` upper CI bound for overall effect size
- `r(tau2)` between-study variance
- `r(I2)` $I^2$ heterogeneity statistic (not for CE model)
- `r(H2)` $H^2$ heterogeneity statistic (not for CE model)
- `r(z)` $z$ statistic for test of significance of overall effect size (when `se()` not specified)
- `r(t)` $t$ statistic for test of significance of overall effect size (when `se()` specified)
- `r(df)` degrees of freedom for $t$ distribution
- `r(p)` $p$-value for test of significance of overall effect size
- `r(Q)` Cochran’s $Q$ heterogeneity test statistic (not for CE model)
- `r(df_Q)` degrees of freedom for heterogeneity test
- `r(p_Q)` $p$-value for heterogeneity test
- `r(Q_b)` Cochran’s $Q$ statistic for test of group differences (for `subgroup()` with one variable)
- `r(df_Q_b)` degrees of freedom for test of group differences
- `r(p_Q_b)` $p$-value for test of group differences
- `r(seadj)` standard error adjustment
- `r(level)` confidence level for CIs
- `r(pi_lb)` lower bound of prediction interval
- `r(pi_ub)` upper bound of prediction interval
- `r(pilevel)` confidence level for prediction interval
- `r(converged)` 1 if converged, 0 otherwise (with iterative random-effects methods)

**Macros**
- `r(model)` meta-analysis model
- `r(method)` meta-analysis estimation method
- `r(subgroupvars)` names of subgroup-analysis variables
- `r(ordervar)` name of order variable used in option `cumulative()`
- `r(byvar)` name of variable used in suboption `by()` within option `cumulative()`
- `r(direction)` ascending or descending
- `r(seadjtype)` type of standard error adjustment

**Matrices**
- `r(esgroup)` ESs and CIs from subgroup analysis
- `r(hetgroup)` heterogeneity summary from subgroup analysis
- `r(diffgroup)` results for tests of group differences from subgroup analysis
- `r(cumul)` results from cumulative meta-analysis
- `r(leaveoneout)` results from leave-one-out meta-analysis

`meta summarize` also creates a system variable, `meta_weight`, which contains study weights. When the `transform()` option is specified, `meta summarize` creates system variables `meta_es_transf`, `meta_cil_transf`, and `meta_ciu_transf`, which contain the transformed effect sizes and lower and upper bounds of the corresponding transformed CIs.

Also see `Stored results` in `[META] meta set` and `Stored results` in `[META] meta esize` for other system variables.
Methods and formulas

Methods and formulas are presented under the following headings:

- Fixed-effects and common-effect methods for combining study estimates
  - Inverse-variance method
  - Mantel–Haenszel method for binary outcomes
  - Peto’s method for odds ratios
- Random-effects methods for combining study estimates
  - Iterative methods
  - Noniterative methods
  - Knapp–Hartung standard-error adjustment
  - Prediction intervals
- Confidence intervals and significance test
- Heterogeneity measures
- Homogeneity test
- Subgroup meta-analysis
  - Fixed-effects model
  - Random-effects model
- Cumulative meta-analysis
  - Leave-one-out meta-analysis

The formulas and methods below are based on Veroniki et al. (2016), Viechtbauer et al. (2015), Borenstein et al. (2009), Schwarzer, Carpenter, and Rucker (2015), Kontopantelis and Reeves (2016), Fisher (2016), and Bradburn, Deeks, and Altman (2016).

Fixed-effects and common-effect methods for combining study estimates

Consider the data from \( K \) independent studies. Let \( \hat{\theta}_j \) be the estimate of the population effect size \( \theta_j \) reported by the \( j \)th study and \( \hat{\sigma}^2_j \) be the corresponding estimate of the within-study variance, which is equal to the squared standard error of \( \hat{\theta}_j \). \( \hat{\theta}_j \) is one of Hedges’s \( g_j \), Cohen’s \( d_j \), \( \ln(\hat{OR}_j) \), \( \ln(\hat{RR}_j) \), and so on, as defined in Methods and formulas of [META] meta esize, or a generic (precomputed) effect size as declared by [META] meta set.

Consider a fixed-effects model (Hedges and Vevea 1998; Rice, Higgins, and Lumley 2018) from Meta-analysis models in [META] Intro,

\[
\hat{\theta}_j = \theta_j + \epsilon_j \quad \epsilon_j \sim N(0, \hat{\sigma}^2_j)
\]

where \( \hat{\sigma}^2_j \)’s are treated as known values that do not require estimation. Under the assumption that \( \theta_1 = \theta_2 = \cdots = \theta_K = \theta \), the above fixed-effects model simplifies to a common-effect model (Hedges 1982; Rosenthal and Rubin 1982):

\[
\hat{\theta}_j = \theta + \epsilon_j \quad \epsilon_j \sim N(0, \hat{\sigma}^2_j)
\]

The estimation methods we describe below are the same for the two models, but the interpretation of the estimates is different; see Comparison between the models and interpretation of their results in [META] Intro. The two models estimate different population parameters. A common-effect model estimates the common effect \( \theta_{\text{pop}} = \theta \), whereas a fixed-effects model estimates a weighted average of the study-specific effects \( \hat{\theta}_j \)’s,

\[
\theta_{\text{pop}} = \text{Ave}(\theta_j) = \frac{\sum_{j=1}^{K} W_j \theta_j}{\sum_{j=1}^{K} W_j}
\]

where \( W_j \)’s represent true, unknown weights, which are defined in Rice, Higgins, and Lumley (2018, eq. 3). For simplicity, in what follows, we will use \( \theta \) to mean \( \theta_{\text{pop}} \).
Inverse-variance method

Under the inverse-variance method, the MLE of $\theta$ is

$$\hat{\theta}_{IV} = \frac{\sum_{j=1}^{K} \hat{\theta}_j / \hat{\sigma}_j^2}{\sum_{j=1}^{K} 1 / \hat{\sigma}_j^2} = \frac{\sum_{j=1}^{K} w_j \hat{\theta}_j}{\sum_{j=1}^{K} w_j}$$

where the weight $w_j = 1 / \hat{\sigma}_j^2$ is used to estimate the true weight $W_j$. The inverse-variance method takes its name from the weights being the reciprocal of the effect-size variances.

The variance estimate of $\hat{\theta}_{IV}$

$$\hat{\text{Var}}(\hat{\theta}_{IV}) = \frac{1}{w}.$$ 

where $w = \sum_{j=1}^{K} w_j$.

Mantel–Haenszel method for binary outcomes

For meta-analysis that compares two binary outcomes, the Mantel–Haenszel method can be used to combine odds ratios (OR), risk ratios (RR), and risk differences (RD) instead of the inverse-variance method. The classical Mantel–Haenszel method (Mantel and Haenszel 1959) is used for OR, and its extension by Greenland and Robins (1985) is used for RR and RD. The Mantel–Haenszel method may be preferable with sparse data (Emerson 1994). This is the default pooling method in meta esize for the effect sizes mentioned above with fixed-effects and common-effect models.

Consider the following $2 \times 2$ table for the $j$th study.

<table>
<thead>
<tr>
<th>group</th>
<th>event</th>
<th>no event</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>treatment</td>
<td>$a_j$</td>
<td>$b_j$</td>
<td>$n_{1j} = a_j + b_j$</td>
</tr>
<tr>
<td>control</td>
<td>$c_j$</td>
<td>$d_j$</td>
<td>$n_{2j} = c_j + d_j$</td>
</tr>
</tbody>
</table>

The sample size for the $j$th study is denoted by $n_j = n_{1j} + n_{2j}$.

For the overall risk difference, the formula is

$$\hat{\theta}_{MH} = \frac{\sum_{j=1}^{K} w_j^{(MH)} \times \hat{\theta}_j}{\sum_{j=1}^{K} w_j^{(MH)}}$$

where $\hat{\theta}_j$ is $\hat{\text{RD}}$ from the $j$th study.

Unlike the inverse-variance method, with log odds-ratios and log risk-ratios, the Mantel–Haenszel method combines the individual effect sizes in the original metric and then takes the log to obtain the final overall log odds-ratio or log risk-ratio estimate,

$$\hat{\theta}_{MH} \ = \ \ln \left\{ \frac{\sum_{j=1}^{K} w_j^{(MH)} \times \exp(\hat{\theta}_j)}{\sum_{j=1}^{K} w_j^{(MH)}} \right\}$$

where $\hat{\theta}_j$ is $\ln(\hat{\text{OR}})$ or $\ln(\hat{\text{RR}})$ from the $j$th study.
The MH weights are defined as follows. In the formula for the overall risk difference, the weight assigned to each study is 
\[ w_{j}^{(MH)} = \frac{n_{1j}n_{2j}}{n_{j}} \]
For the overall log risk-ratio, the \( j \)th weight is given by 
\[ w_{j}^{(MH)} = \frac{n_{1j}c_{j}}{n_{j}} \]
And for the overall log odds-ratio, the \( j \)th weight is given by 
\[ w_{j}^{(MH)} = \frac{b_{j}c_{j}}{n_{j}} \]

An estimator of the variance of the overall risk difference \( \hat{\theta}_{MH} = \hat{\text{RD}}_{MH} \) (Greenland and Robins 1985) is 
\[ \hat{\text{Var}}(\hat{\text{RD}}_{MH}) = \sum_{j=1}^{K} \left( \frac{a_{j}b_{j}n_{2j}^{3} + c_{j}d_{j}n_{1j}^{3}}{n_{1j}n_{2j}^{2}} \right) / n_{j}^{2} \]
\[ \left( \sum_{j=1}^{K} n_{1j}n_{2j} / n_{j} \right)^{2} \]
An estimator of the variance of the overall log risk-ratio \( \hat{\theta}_{MH} = \ln(\hat{\text{RR}}_{MH}) \) (Greenland and Robins 1985) is 
\[ \hat{\text{Var}}\{ \ln(\hat{\text{RR}}_{MH}) \} = \sum_{j=1}^{K} \left( \frac{n_{1j}n_{2j} (a_{j} + c_{j}) - a_{j}c_{j}n_{j}}{n_{j}^{2}} \right) \]
\[ \left( \sum_{j=1}^{K} a_{j}n_{2j} / n_{j} \right) \times \left( \sum_{j=1}^{K} c_{j}n_{1j} / n_{j} \right) \]
And an estimator of the variance of the overall log odds-ratio \( \hat{\theta}_{MH} = \ln(\hat{\text{OR}}_{MH}) \) (Robins, Breslow, and Greenland 1986a; Robins, Greenland, and Breslow 1986b) is 
\[ \hat{\text{Var}}\{ \ln(\hat{\text{OR}}_{MH}) \} = \frac{\sum_{j=1}^{K} P_{j}R_{j}}{2\left( \sum_{j=1}^{K} R_{j} \right)^{2}} + \frac{\sum_{j=1}^{K} (P_{j}S_{j} + Q_{j}R_{j})}{2\sum_{j=1}^{K} R_{j} \sum_{j=1}^{K} S_{j}} + \frac{\sum_{j=1}^{K} Q_{j}S_{j}}{2\left( \sum_{j=1}^{K} S_{j} \right)^{2}} \]
where 
\[ P_{j} = \frac{a_{j} + d_{j}}{n_{j}}, \quad Q_{j} = \frac{b_{j} + c_{j}}{n_{j}}, \quad R_{j} = \frac{a_{j}d_{j}}{n_{j}}, \quad \text{and} \quad S_{j} = \frac{b_{j}c_{j}}{n_{j}} \]

Greenland and Robins (1985) and Robins, Breslow, and Greenland (1986a) demonstrate consistency of all the above variance estimators in the two cases they call a sparse-data limiting model, in which the number of \( 2 \times 2 \) tables (studies) increases but the cell sizes remain fixed, and a large-strata limiting model, in which the number of studies remains fixed but individual cell sizes increase.

**Peto’s method for odds ratios**

An alternative to the Mantel–Haenszel method for combining odds ratios is the Peto’s method (Peto et al. 1977; Yusuf et al. 1985). It is based on the inverse-variance method but uses an alternate way to compute the odds ratios (and consequently the log odds-ratio).
Let \( \ln(\hat{\text{OR}}_{j}^{\text{Peto}}) \) be Peto’s log odds-ratio for the \( j \)th study as defined in *Odds ratio* in [META] *meta esize*. Then, Peto’s overall log odds-ratio is defined following the inverse-variance method as follows,

\[
\hat{\theta}_{\text{Peto}} = \ln\left(\hat{\text{OR}}_{j}^{\text{Peto}}\right) = \frac{\sum_{j=1}^{K} w_{j} \ln(\hat{\text{OR}}_{j}^{\text{Peto}})}{\sum_{j=1}^{K} w_{j}}
\]

where \( w_{j} = 1/\hat{\sigma}_{j}^{2} = \text{Var}(a_{j}) \) and \( \text{Var}(a_{j}) \) is as defined in *Methods and formulas* of [META] *meta esize* of [META] *meta esize*.

The variance estimate is

\[
\hat{\text{Var}}\left\{ \ln(\hat{\text{OR}}_{j}^{\text{Peto}}) \right\} = \frac{1}{\sum_{j=1}^{K} w_{j}}
\]

### Random-effects methods for combining study estimates

Suppose that the observed study-specific effect sizes represent a random sample from a population of effect sizes that is normally distributed with mean \( \theta \) and variance \( \tau^{2} \).

Consider a random-effects model (Hedges 1983; DerSimonian and Laird 1986) from *Meta-analysis models* in [META] *Intro*,

\[
\hat{\theta}_{j} = \theta_{j} + \epsilon_{j} = \theta + u_{j} + \epsilon_{j}
\]

where \( \epsilon_{j} \) and \( u_{j} \) are assumed to be independent with \( \epsilon_{j} \sim N\left(0, \hat{\sigma}_{j}^{2}\right) \) and \( u_{j} \sim N\left(0, \tau^{2}\right) \).

The overall effect \( E(\hat{\theta}_{j}) = \theta \) is estimated as the weighted average,

\[
\hat{\theta}^{*} = \frac{\sum_{j=1}^{K} w_{j}^{*} \hat{\theta}_{j}}{\sum_{j=1}^{K} w_{j}^{*}}
\]

(1)

where \( w_{j}^{*} = 1/\left(\hat{\sigma}_{j}^{2} + \tilde{\tau}^{2}\right) \). The variance of \( \hat{\theta}^{*} \) is estimated by

\[
\hat{\text{Var}}(\hat{\theta}^{*}) = \frac{1}{w^{*}}
\]

where \( w^{*} = \sum_{j=1}^{K} w_{j}^{*} \).

Different estimators of the between-study variance, \( \tau^{2} \), lead to different estimators of \( \theta \). *meta summarize* supports seven estimation methods of \( \tau^{2} \). Three methods are iterative: the maximum likelihood (ML) estimator (Hardy and Thompson 1996); the restricted maximum-likelihood (REML) estimator (Raudenbush 2009); and the empirical Bayes (EB) estimator (Morris 1983; Berkey et al. 1995), also known as the Paule–Mandel estimator (Paule and Mandel 1982). Four methods are noniterative (have a closed-form expression): DerSimonian–Laird (DL) estimator (DerSimonian and Laird 1986); Hedges estimator (HE) (Hedges 1983; Hedges and Olkin 1985), also known as the Cochran estimator or variance-component estimator; Hunter–Schmidt (HS) estimator (Schmidt and Hunter 2015); and Sidik–Jonkman (SJ) estimator (Sidik and Jonkman 2005).

The formulas for and properties of these estimators have been discussed at length in Veroniki et al. (2016). Expressions for these estimators are given in the more general context of meta-regression in *Methods and formulas* of [META] *meta regress*. Below, we provide the simplified expressions when no covariates (moderators) are included in the regression model. The simplified expressions were obtained by replacing the \( X \) matrix with \( K \times 1 \) column vector of 1s.
Iterative methods

The ML method (Hardy and Thompson 1996; Thompson and Sharp 1999) computes the MLE of \(\tau^2\) by maximizing the following log-likelihood function,

\[
\ln L_{\text{ML}} (\tau^2) = -\frac{K}{2} \ln(2\pi) - \frac{1}{2} \sum_{j=1}^{K} \ln (\hat{\sigma}_j^2 + \tau^2) - \frac{1}{2} \sum_{j=1}^{K} \frac{(\hat{\theta}_j - \hat{\theta}^*)^2}{\hat{\sigma}_j^2 + \tau^2}
\]

with respect to \(\tau^2\), where \(\hat{\theta}^*\) is defined in (1) and is based on the current value of \(\hat{\tau}^2\).

The ML method is asymptotically efficient but may produce biased results in small samples. The REML method estimates \(\tau^2\) by accounting for the uncertainty in the estimation of \(\theta\), which leads to nearly an unbiased estimate of \(\tau^2\).

The REML log-likelihood function is

\[
\ln L_{\text{REML}} (\tau^2) = \ln L_{\text{ML}} (\tau^2) - \frac{1}{2} \ln \left( \sum_{j=1}^{K} (\hat{\sigma}_j^2 + \tau^2) \right)^{-1} + \frac{\ln(2\pi)}{2}
\]

The EB estimator and a description of the iterative process for each estimator in this section is presented in the Methods and formulas of [META] meta regress.

Noniterative methods

The methods in this section do not make any assumptions about the distribution of the random effects. They also do not require any iteration.

The most popular noniterative estimation method is the DL method. This is a method of moment estimator for \(\tau^2\), and it is defined as follows,

\[
\hat{\tau}^2_{\text{DL}} = \frac{Q - (K - 1)}{\sum_{j=1}^{K} w_j - \sum_{j=1}^{K} w_j^2 / \sum_{j=1}^{K} w_j}
\]

where \(Q = \sum_{j=1}^{K} w_j (\hat{\theta}_j - \hat{\theta}_{IV})^2\) and \(w_j = 1/\hat{\sigma}_j^2\).

Because \(\hat{\tau}^2_{\text{DL}}\) is negative when \(Q < K - 1\), it is truncated at 0 in practice, and thus \(\max(0, \hat{\tau}^2_{\text{DL}})\) is used to estimate the between-study variance:

\[
\hat{\tau}^2_{\text{DL}} = \max \left\{ 0, \frac{\sum_{j=1}^{K} w_j (\hat{\theta}_j - \hat{\theta}_{IV})^2 - (K - 1)}{\sum_{j=1}^{K} w_j - \sum_{j=1}^{K} w_j^2 / \sum_{j=1}^{K} w_j} \right\}
\]

The HE estimator is another method of moment estimator defined as follows,

\[
\hat{\tau}^2_{\text{HE}} = \max \left\{ 0, \frac{1}{K - 1} \sum_{j=1}^{K} (\hat{\theta}_j - \bar{\theta})^2 - \frac{1}{K} \sum_{j=1}^{K} \hat{\sigma}_j^2 \right\}
\]

where \(\bar{\theta} = (\sum_{j=1}^{K} \hat{\theta}_j)/K\).
The HS estimator is given by

\[ \hat{\tau}^2_{\text{HS}} = \max \left\{ 0, \frac{Q - K}{\sum_{j=1}^{K} w_j} \right\} \]

For the SJ estimator, consider an initial estimate of \( \tau^2 \), given by

\[ \hat{\tau}_0^2 = \sum_{j=1}^{K} \left( \hat{\theta}_j - \bar{\theta} \right)^2 \]

Then, the estimator is defined as

\[ \hat{\tau}^2_{\text{SJ}} = \frac{\sum_{j=1}^{K} w_{SJ,j} \left( \hat{\theta}_j - \hat{\theta}_{\text{SJ}} \right)^2}{K - 1} \]

where \( w_{SJ,j}^2 = \frac{\hat{\tau}_0^2}{\hat{\sigma}_j^2 + \hat{\tau}_0^2} \) and \( \hat{\theta}_{\text{SJ}} = \sum_{j=1}^{K} w_{SJ,j} \hat{\theta}_j / \sum_{j=1}^{K} w_{SJ,j} \).

**Knapp–Hartung standard-error adjustment**

Hartung and Knapp (2001a) and Sidik and Jonkman (2002) proposed an adjustment to the variance of \( \hat{\theta}^* \) to account for the uncertainty in estimating \( \tau^2 \), which is used in the expression for weights. They proposed to multiply \( \hat{\text{Var}}(\hat{\theta}^*) = 1/w^* \) by the following quadratic form,

\[ q_{\text{KH}} = \frac{1}{K - 1} \sum_{j=1}^{K} w_j^* \left( \hat{\theta}_j - \hat{\theta}^* \right)^2 \]

or by \( \max (1, q_{\text{KH}}) \).

The variance estimator for \( \hat{\theta}^* \) can then be defined as

\[ \hat{\text{Var}}_{\text{HK}} (\hat{\theta}^*) = \begin{cases} q_{\text{KH}} \times 1/w^* & \text{with option se(khartung)} \\ \max (1, q_{\text{KH}}) \times 1/w^* & \text{with option se(khartung, truncated)} \end{cases} \]

Hartung (1999) established that the statistic

\[ \frac{\hat{\theta}^* - \theta}{\sqrt{\hat{\text{Var}}_{\text{HK}} (\hat{\theta}^*)}} \]

has a Student’s \( t \) distribution with \( K - 1 \) degrees of freedom.
Correspondingly, the \( (1 - \alpha) \times 100\% \) CI for \( \theta \) using the Knapp–Hartung standard error is

\[
\hat{\theta}^* \pm t_{K-1,1-\alpha/2} \sqrt{\text{Var}_{HK}(\hat{\theta}^*)}
\]

where \( t_{K-1,1-\alpha/2} \) denotes the \( 1 - \alpha/2 \) quantile of the Student’s \( t \) distribution with \( K - 1 \) degrees of freedom.

The test statistic for the significance test of an overall effect, \( H_0: \theta = 0 \), is

\[
\frac{\hat{\theta}^*}{\sqrt{\text{Var}_{HK}(\hat{\theta}^*)}}
\]

and has the Student’s \( t \) distribution with \( K - 1 \) degrees of freedom.

Also see Sidik and Jonkman (2002, 2003) and Cornell et al. (2014) for more discussion about the Knapp–Hartung adjustment.

**Prediction intervals**

In a random-effects model, you can compute a prediction interval (Higgins, Thompson, and Spiegelhalter 2009) that estimates plausible ranges for \( \theta \) in a future study. Compared with the CI, a prediction interval incorporates the uncertainty in estimating \( \tau^2 \) in the computation.

A \( (1 - \alpha) \times 100\% \) prediction interval is defined as

\[
\hat{\theta}^* \pm t_{K-2,1-\alpha/2} \sqrt{\text{Var}(\hat{\theta}^*) + \hat{\tau}^2}
\]

where \( t_{K-1,1-\alpha/2} \) denotes the \( 1 - \alpha/2 \) quantile of the Student’s \( t \) distribution with \( K - 2 \) degrees of freedom. This prediction interval may be specified with the `predinterval()` option.

**Confidence intervals and significance test**

Let \( \hat{\theta} \) be any of the estimators considered in the previous sections such as \( \hat{\theta}_{IV} \) or \( \hat{\theta}^* \). The \( (1 - \alpha) \times 100\% \) confidence interval for \( \theta \) is

\[
\hat{\theta} \pm z_{1-\alpha/2} \sqrt{\text{Var}(\hat{\theta})}
\]

where \( z_{1-\alpha/2} \) is the \( (1 - \alpha/2) \)th quantile of the standard normal distribution.

We reject the hypothesis of no treatment effect \( H_0: \theta = 0 \) at level \( \alpha \), if

\[
\frac{\hat{\theta}}{\sqrt{\text{Var}(\hat{\theta})}} > z_{1-\alpha/2}
\]

If the `tdistribution` option is specified, the \( z_{1-\alpha/2} \) critical value is replaced with the \( t_{K-1,1-\alpha/2} \) critical value in the above formulas.
Heterogeneity measures

The homogeneity test can be used to test whether the study-specific effects are the same; see Homogeneity test. But with a small number of studies, this test may have low power (Hedges and Pigott 2001). Also, it does not provide an estimate of the magnitude of the between-study heterogeneity. Some authors (for example, Higgins and Thompson [2002] and Higgins et al. [2003]) suggest examining the heterogeneity statistics rather than relying solely on the homogeneity test.

Higgins and Thompson (2002) proposed two heterogeneity measures: $I^2$ and $H^2$. We define them separately for random-effects and fixed-effects models.

For a random-effects model, the two heterogeneity measures are defined as follows:

$$I^2 = \frac{\hat{\tau}^2}{\hat{\tau}^2 + s^2} \times 100\%$$

and

$$H^2 = \frac{\hat{\tau}^2 + s^2}{s^2}$$

where

$$s^2 = \frac{K - 1}{\sum_{j=1}^{K} w_j - \sum_{j=1}^{K} w_j^2 / \sum_{j=1}^{K} w_j}$$

is the within-study variance and $\hat{\tau}^2$ is an estimator of the between-study variance. The values of $I^2$ and $H^2$ will vary depending on which estimator of $\hat{\tau}^2$ is specified in the random() option.

For a fixed-effects model, the expressions for $I^2$ and $H^2$ are given by

$$I^2 = \left\{ \frac{Q - (K - 1)}{Q} \right\} \times 100\%$$

and

$$H^2 = \frac{Q}{K - 1}$$

where $Q$ is defined in Homogeneity test.

The formulas above for $I^2$ and $H^2$ are equivalent to the corresponding formulas (2) and (3), when the DL method is used to estimate $\tau^2$. $I^2$ is negative when $Q < (K - 1)$ and is thus reset to zero in that case.

Homogeneity test

Consider a test of $H_0$: $\theta_1 = \theta_2 = \cdots = \theta_K = \theta$, known as the homogeneity test, that evaluates whether the effect sizes are the same across the studies. It uses the following test statistic,

$$Q = \sum_{j=1}^{K} w_j \left( \hat{\theta}_j - \hat{\theta} \right)^2 = \sum_{j=1}^{K} w_j \hat{\theta}_j^2 - \left( \frac{\sum_{j=1}^{K} w_j \hat{\theta}_j}{\sum_{j=1}^{K} w_j} \right)^2$$

where $w_j = 1/\hat{\sigma}_j^2$, and $\hat{\theta}_j$ and $\hat{\theta}$ depend on the type of the effect size chosen.

Under the null hypothesis of homogeneity, $Q$ follows a $\chi^2$ distribution with $K - 1$ degrees of freedom.
Hedges and Pigott (2001) showed that the test has low power when the number of studies \((K)\) is small, which is typical in meta-analysis. This means that the null hypothesis of homogeneity is not rejected as often as it should be. Thus, for the homogeneity test, the meta-analysis literature (for example, Petitti [2001]; Berman and Parker [2002]; Sutton and Higgins [2008]) suggests using the significance level \(\alpha = 0.1\) instead of the conventional \(\alpha = 0.05\).

The homogeneity test checks for the potential presence of heterogeneity but does not estimate the magnitude of the heterogeneity. Thus, many authors (for example, Higgins and Thompson [2002]; Higgins et al. [2003]) suggest exploring the heterogeneity statistics rather than solely relying on the test. See Heterogeneity measures.

**Subgroup meta-analysis**

When the subgroup(varname) option is specified, we assume that the \(K\) studies are partitioned into \(L\) subgroups defined by varname. Estimates of the overall effect size and their corresponding standard errors are calculated for each of the \(L\) subgroups.

Let \(\hat{\theta}_{jl}\) be the effect-size estimate from study \(j\) within subgroup \(l\) and \(\hat{\sigma}_{jl}^2\) be the corresponding variance, where \(l = 1, 2, \ldots, L\) and \(j = 1, 2, \ldots, K_l\).

Below, we describe the formulas separately for fixed-effects and random-effects models. The formulas for the common-effect model are the same as for the fixed-effects model. When you specify a common-effect model with subgroup analysis, this model is assumed within each subgroup \(l = 1, 2, \ldots, L\), but not for the entire sample of studies.

**Fixed-effects model**

In what follows, we assume the inverse-variance method, but the same principles apply to the Mantel–Haenszel method.

In subgroup analysis, a fixed-effects model may be formulated as

\[
\hat{\theta}_{jl} = \theta_{jl} + \epsilon_{jl}, \quad \epsilon_{jl} \sim N(0, \hat{\sigma}_{jl}^2)
\]

For the \(l\)th group, \(\hat{\theta}_{IV,l}\) is a weighted average of the effect sizes \(\hat{\theta}_{jl}\) with weights \(w_{jl} = 1/\hat{\sigma}_{jl}^2\):

\[
\hat{\theta}_{IV,l} = \frac{\sum_{j=1}^{K_l} w_{jl} \hat{\theta}_{jl}}{\sum_{j=1}^{K_l} w_{jl}}
\]

The variance estimate of \(\hat{\theta}_{IV,l}\) is

\[
\hat{\text{Var}}(\hat{\theta}_{IV,l}) = \frac{1}{w_{.l}}
\]

where \(w_{.l} = \sum_{j=1}^{K_l} w_{jl}\).

Other meta-analytic quantities such as \(I_l^2\) and \(Q_l\) may also be computed for the \(l\)th subgroup just as we described in the previous sections.
The Cochran’s $Q$ statistic can be extended to test for differences between the $L$ subgroups:

$$Q_b = \sum_{l=1}^{L} w_{.l} \left( \hat{\theta}_{IV,l} - \frac{\sum_{l=1}^{L} w_{.l} \hat{\theta}_{IV,l}}{\sum_{l=1}^{L} w_{.l}} \right)^2$$

The subscript $b$ in $Q_b$ stands for “between” to emphasize that $Q_b$ tests for “between-group” differences.

Under the null hypothesis of homogeneity between the subgroups ($\theta_{.1} = \theta_{.2} = \cdots = \theta_{.L} = \theta$), the statistic $Q_b$ has a $\chi^2$ distribution with $L - 1$ degrees of freedom.

### Random-effects model

Consider a random-effects model with $L$ subgroups and separate between-study variances $\tau^2_l$:

$$\hat{\theta}_{jl} = \theta_{.l} + u_{jl} + \epsilon_{jl} \quad \epsilon_{jl} \sim N(0, \hat{\sigma}^2_{jl}) \quad u_{jl} \sim N(0, \tau^2_l)$$

The formulas for the random-effects model are the same as for the above fixed-effects model, except we replace the weights with the random-effects weights.

The estimate, $\hat{\theta}^*_{l}$, and its variance in the $l$th group are

$$\hat{\theta}^*_{l} = \frac{\sum_{j=1}^{K_l} w^*_{jl} \hat{\theta}_{jl}}{\sum_{j=1}^{K_l} w^*_{jl}}$$

$$\widehat{\text{Var}} \left( \hat{\theta}^*_{l} \right) = \frac{1}{w^*_{.l}}$$

where $w^*_{jl} = 1/(\hat{\sigma}^2_{jl} + \hat{\tau}^2_l)$ and $w^*_{.l} = \sum_{j=1}^{K_l} w^*_{jl}$.

The Cochran’s statistic for testing differences between the $L$ subgroups is defined as

$$Q^*_b = \sum_{l=1}^{L} w^*_{.l} \left( \hat{\theta}^*_{l} - \frac{\sum_{l=1}^{L} w^*_{.l} \hat{\theta}^*_{l}}{\sum_{l=1}^{L} w^*_{.l}} \right)^2$$

Under the null hypothesis of homogeneity between the subgroups ($\theta_{.1} = \theta_{.2} = \cdots = \theta_{.L} = \theta$), $Q^*_b$ has a $\chi^2$ distribution with $L - 1$ degrees of freedom.

Also see Borenstein et al. (2009, chap. 19) and Schwarzer, Carpenter, and Rücker (2015).

### Cumulative meta-analysis

To perform CMA, we first sort the studies in ascending order according to the values of the variable specified in the `cumulative()` option. If suboption `descending` is specified within the `cumulative()` option, the order is reversed. Mathematically, this corresponds to sorting the pairs $(\hat{\theta}_j, \hat{\sigma}^2_j)$ in the specified order. Let $(\hat{\theta}^*_{j}, \hat{\sigma}^2_{j}^{*})$ denote the sorted pairs.
CMA estimates $K$ overall effect sizes $\hat{\theta}_j^c$'s as follows,

$$\hat{\theta}_1^c = \hat{\theta}_1^s$$
$$\hat{\theta}_2^c = \text{MA}(\hat{\theta}_1^s, \hat{\theta}_2^s)$$
$$\hat{\theta}_3^c = \text{MA}(\hat{\theta}_1^s, \hat{\theta}_2^s, \hat{\theta}_3^s)$$
$$\vdots$$
$$\hat{\theta}_j^c = \text{MA}(\hat{\theta}_1^s, \hat{\theta}_2^s, \hat{\theta}_3^s, \ldots, \hat{\theta}_j^s)$$
$$\vdots$$
$$\hat{\theta}_K^c = \text{MA}(\hat{\theta}_1^s, \hat{\theta}_2^s, \hat{\theta}_3^s, \ldots, \hat{\theta}_K^s)$$

where $\text{MA}(\hat{\theta}_1^s, \hat{\theta}_2^s, \hat{\theta}_3^s, \ldots, \hat{\theta}_j^s)$ denotes a meta-analysis applied to the sorted studies 1 through $j$. Note that the meta-analysis also depends on the values $\hat{\sigma}_j^2$s but we omitted them from $\text{MA}()$ for notational convenience.

If suboption by(byvar) is specified within the cumulative() option, the above procedure is repeated for each subgroup defined by variable byvar.

**Leave-one-out meta-analysis**

Leave-one-out meta-analysis estimates $K$ overall effect sizes $\hat{\theta}_{-j}$'s as follows,

$$\hat{\theta}_{-1} = \text{MA}(\hat{\theta}_2, \hat{\theta}_3, \ldots, \hat{\theta}_K)$$
$$\hat{\theta}_{-2} = \text{MA}(\hat{\theta}_1, \hat{\theta}_3, \ldots, \hat{\theta}_K)$$
$$\vdots$$
$$\hat{\theta}_{-j} = \text{MA}(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{j-1}, \hat{\theta}_{j+1}, \ldots, \hat{\theta}_K)$$
$$\vdots$$
$$\hat{\theta}_{-K} = \text{MA}(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \ldots, \hat{\theta}_{K-1})$$

where $\text{MA}(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{j-1}, \hat{\theta}_{j+1}, \ldots, \hat{\theta}_K)$ denotes a meta-analysis applied to all the studies except the $j$th study. Note that the meta-analysis also depends on the values $\hat{\sigma}_j^2$, but we omitted them from $\text{MA}()$ for notational convenience.
References


Also see

- [META] meta data — Declare meta-analysis data
- [META] meta forestplot — Forest plots
- [META] meta galbraithplot — Galbraith plots
- [META] meta regress — Meta-analysis regression
- [META] meta — Introduction to meta
- [META] Glossary
- [META] Intro — Introduction to meta-analysis