Title

> **meta esize —** Compute effect sizes and declare meta-analysis data

Description

meta esize computes effect sizes from study summary data and uses the results to declare the data in memory to be meta data, informing Stata of key variables and their roles in a meta-analysis. It computes various effect sizes and their respective standard errors for two-group comparisons of continuous and binary outcomes and for estimating a single proportion of a binary outcome. It then uses the computed effect sizes and standard errors to declare the data in memory to be meta data. If you do not have the summary data from individual studies and, instead, you have precalculated effect sizes, you can use meta set to declare your meta-analysis data. You must use meta esize or meta set to perform univariate meta-analysis using the meta command; see [META] **meta data**.

If you need to update some of the meta settings after the data declaration, see [META] **meta update**. To display current meta settings, use meta query; see [META] **meta update**.

Quick start

Compute Hedges's g standardized mean differences and their standard errors from variables nt (sample size in treatment group), meant (mean of treatment group), sdt (standard deviation in treatment group), and their counterparts in the control group: nc, meanc, and sdc

    meta esize nt meant sdt nc meanc sdc

Same as above, but compute Cohen's d instead of the default Hedges's g, and use the DerSimonian–Laird estimation method instead of the default REML method

    meta esize nt meant sdt nc meanc sdc, esize(cohend) random(dlaird)

Compute log odds-ratios and their standard errors from variables nst (number of successes in treatment group), nft (number of failures in treatment group), and their respective counterparts in control group: nsc and nfc

    meta esize nst nft nsc nfc

Same as above, but compute the log risk-ratios instead of the default log odds-ratios

    meta esize nst nft nsc nfc, esize(lnrratio)

Same as above, but request a common-effect meta-analysis

    meta esize nst nft nsc nfc, esize(lnrratio) common

Compute transformed proportions using the default Freeman–Tukey double-arcsine transformation and their standard errors from variables ns (number of successes) and n (study sample size)

    meta esize ns n

Same as above, but compute the logit-transformed proportions instead of the default Freeman–Tukey-transformed proportions

    meta esize ns n, esize(logitprop)

Menu

Statistics > Meta-analysis

Syntax

Compute and declare effect sizes for two-group comparison of continuous outcomes

>   meta <u>esize</u> *n1 mean1 sd1 n2 mean2 sd2* [*if*] [*in*] [, *options_continuous options*]

Compute and declare effect sizes for two-group comparison of binary outcomes

>   meta <u>esize</u> *n11 n12 n21 n22* [*if*] [*in*] [, *options_binary options*]

Compute and declare effect sizes for estimating a single proportion (prevalence)

>   meta <u>esize</u> *ns n* [*if*] [*in*] [, *options_proportion options*]

Variables *n1*, *mean1*, and *sd1* contain sample sizes, means, and standard deviations from individual studies for group 1 (treatment), and variables *n2*, *mean2*, and *sd2* contain the respective summaries for group 2 (control).

Variables *n11* and *n12* contain numbers of successes and numbers of failures from individual studies for group 1 (treatment), and variables *n21* and *n22* contain the respective numbers for group 2 (control). A single observation defined by variables *n11*, *n12*, *n21*, and *n22* represents a 2×2 table from an individual study. Therefore, variables *n11*, *n12*, *n21*, and *n22* represent a sample of 2×2 tables from all studies. We will thus refer to observations on these variables as 2×2 tables and to values of these variables as cells.

Variables *ns* and *n* contain number of successes and sample sizes from individual studies. Here, the values of the variable *ns* (number of successes) and the values of an implicit "variable" $n - ns$ (number of failures) are referred to as cells.

options_continuous	Description
Main	
<u>esize</u>(*esspeccnt*)	specify effect size for two-group comparison of continuous outcomes to be used in the meta-analysis
Model	
<u>random</u>[(*remethod*)]	random-effects meta-analysis; default is random(reml)
<u>common</u>	common-effect meta-analysis; implies inverse-variance method
<u>fixed</u>	fixed-effects meta-analysis; implies inverse-variance method

options_binary	Description
Main	
<u>esize</u>(*estypebin*)	specify effect size for two-group comparison of binary outcomes to be used in the meta-analysis
Model	
<u>random</u>[(*remethod*)]	random-effects meta-analysis; default is random(reml)
<u>common</u>[(*cefemethod*)]	common-effect meta-analysis
<u>fixed</u>[(*cefemethod*)]	fixed-effects meta-analysis
Options	
<u>zerocells</u>(*zcspec*)	adjust for zero cells during computation; default is to add 0.5 to all cells of those 2×2 tables that contain zero cells

options_proportion	Description
Main	
<u>esize</u>(*estypeprop*)	specify effect size for estimating a single proportion to be used in the meta-analysis
Model	
<u>random</u>[(*remethod*)]	random-effects meta-analysis; default is random(reml)
common	common-effect meta-analysis; implies inverse-variance method
fixed	fixed-effects meta-analysis; implies inverse-variance method
Options	
zerocells(*zcspec*)	adjust for zero cells during computation; default is to add 0.5 to all cells of studies with zero successes or failures

options	Description
Options	
studylabel(*varname*)	variable to be used to label studies in all meta-analysis output
eslabel(*string*)	effect-size label to be used in all meta-analysis output; default is eslabel(Effect size)
<u>level</u>(#)	confidence level for all subsequent meta-analysis commands
[no]metashow	display or suppress meta settings with other meta commands

The syntax of *esspeccnt* is

> *estypecnt* [, *esopts*]

estypecnt	Description
<u>hedgesg</u>	Hedges's g standardized mean difference; the default
<u>cohend</u>	Cohen's d standardized mean difference
<u>glassdelta2</u>	Glass's Δ mean difference standardized by group 2 (control) standard deviation; more common than glassdelta1
<u>glassdelta1</u>	Glass's Δ mean difference standardized by group 1 (treatment) standard deviation
<u>mdiff</u>	(unstandardized) mean difference

estypebin	Description
<u>lnoratio</u>	log odds-ratio; the default
<u>lnrratio</u>	log risk-ratio (also known as log rate-ratio and log relative-risk)
<u>rdiff</u>	risk difference
<u>lnorpeto</u>	Peto's log odds-ratio

estypeprop	Description
<u>ftukey</u>prop	Freeman–Tukey-transformed proportion; the default
<u>logit</u>prop	logit-transformed proportion
<u>proportion</u>	untransformed (raw) proportion

remethod	Description
reml	restricted maximum likelihood; the default
mle	maximum likelihood
<u>ebayes</u>	empirical Bayes
<u>dlaird</u>	DerSimonian–Laird
<u>sjonkman</u>	Sidik–Jonkman
<u>hedges</u>	Hedges
<u>hschmidt</u>	Hunter–Schmidt

cefemethod	Description
<u>mh</u>aenszel	Mantel–Haenszel
<u>invariance</u>	inverse variance
<u>ivariance</u>	synonym for invvariance

Options

$\boxed{\text{Main}}$

esize(*esspec*) specifies the effect size to be used in the meta-analysis. For a two-group comparison of continuous outcomes, *esspec* is *[estypecnt](#)* [, *[esopts](#)*]. For binary outcomes, *esspec* is *[estypebin](#)* for a two-group comparison or *[estypeprop](#)* for estimating a single proportion (prevalence).

For a two-group comparison of continuous outcomes, *estypecnt* is one of the following: hedgesg, cohend, glassdelta2, glassdelta1, or mdiff. Below, we describe each type with its specific options, *esopts*.

hedgesg [, exact holkinse] computes the effect size as the Hedges's g (1981) standardized mean difference. This is the default. For consistency with meta-analysis literature, hedgesg uses an approximation to compute g rather than the exact computation (see *Methods and formulas*), as provided by [esize](#)'s option hedgesg. You can use the exact suboption to match the results from esize (see [R] **esize**).

cohend [, holkinse] computes the effect size as the Cohen's d (1969, 1988) standardized mean difference.

glassdelta2 computes the effect size as the Glass's Δ standardized mean difference, where the standardization uses the standard deviation of the group 2 (control group). glassdelta2 is more common in practice than glassdelta1.

glassdelta1 computes the effect size as the Glass's Δ standardized mean difference, where the standardization uses the standard deviation of the group 1 (treatment group). glassdelta2 is more common in practice than glassdelta1.

mdiff [, unequal] computes the effect size as the unstandardized or raw mean difference.

esopts are `exact`, `holkinse`, and `unequal`.

> `exact` specifies that the exact computation be used for the bias-correction factor in Hedges's g instead of an approximation used by default.

> `holkinse` specifies that the standard error of Hedges's g and Cohen's d be computed as described in Hedges and Olkin (1985). This is another approximation to the standard error of these effect sizes sometimes used in practice.

> `unequal` specifies that the computation of the standard error of the mean difference (`esize(mdiff)`) assume unequal group variances.

For a two-group comparison of binary outcomes, *estypebin* is one of the following: `lnoratio`, `lnrratio`, `rdiff`, or `lnorpeto`.

> `lnoratio` specifies that the effect size be the log odds-ratio. This is the default.

> `lnrratio` specifies that the effect size be the log risk-ratio, also known as a log relative-risk and a log risk-rate.

> `rdiff` specifies that the effect size be the risk difference.

> `lnorpeto` specifies that the effect size be the log odds-ratio as defined by Peto et al. (1977). This effect size is preferable with rare events.

For estimating a proportion from one-sample binary or prevalence data, *estypeprop* is one of the following: `ftukeyprop`, `logitprop`, or `proportion`.

> `ftukeyprop` specifies that the effect size be the Freeman–Tukey-transformed proportion (Freeman and Tukey 1950). This is the default. The Freeman–Tukey transformation is a variance-stabilizing transformation and is preferable when the estimated proportions are close to 0 or 1. This effect size does not require a zero-cell adjustment (continuity correction) for studies with zero successes or failures.

> `logitprop` specifies that the effect size be the logit-transformed proportion. When a study proportion is close to 0 or 1, the estimated variance of this effect size is very large, and thus the study is assigned an artificially small weight in the meta-analysis.

> `proportion` specifies that the effect size be the raw (untransformed) proportion. When a study proportion is close to 0 or 1, its estimated variance is very small, and thus the study is assigned an artificially large weight in the meta-analysis. Moreover, the study confidence limits may fall outside the $[0, 1]$ range.

For effect sizes in the log metric such as log odds-ratios, the results by default are displayed in the log metric. You can use *eform_option* to obtain exponentiated results such as odds ratios. For effect sizes `ftukeyprop` and `logitprop`, the results by default are displayed in the respective Freeman–Tukey and logit metrics. You can use options `transform()` and `proportion` to report results as proportions.

```
Model
```

Options `random()`, `common()`, and `fixed()` declare the meta-analysis model globally throughout the entire meta-analysis; see *Declaring a meta-analysis model* in [META] **meta data**. In other words, once you set your meta-analysis model using `meta esize`, all subsequent `meta` commands will assume that same model. You can update the declared model by using `meta update` or change it temporarily by specifying the corresponding option with the `meta` commands. Options `random()`, `common()`, and `fixed()` may not be combined. If these options are omitted, `random(reml)` is assumed; see *Default meta-analysis model and method* 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 τ^2. *remethod* is one of reml, mle, ebayes, dlaird, sjonkman, hedges, or hschmidt. random is a synonym for random(reml). Below, we provide a short description for each method based on Veroniki et al. (2016). Also see *Declaring a meta-analysis estimation method* in [META] **meta data**.

  reml, the default, specifies that the REML method (Raudenbush 2009) be used to estimate τ^2. This method produces an unbiased, nonnegative estimate of the between-study variance and is commonly used in practice. Method reml requires iteration.

  mle specifies that the ML method (Hardy and Thompson 1996) be used to estimate τ^2. It produces a nonnegative estimate of the between-study variance. With a few studies or small studies, this method may produce biased estimates. With many studies, the ML method is more efficient than the REML method. Method mle requires iteration.

  ebayes specifies that the empirical Bayes estimator (Berkey et al. 1995), also known as the Paule–Mandel estimator (Paule and Mandel 1982), be used to estimate τ^2. From simulations, this method, in general, tends to be less biased than other random-effects methods, but it is also less efficient than reml or dlaird. Method ebayes produces a nonnegative estimate of τ^2 and requires iteration.

  dlaird specifies that the DerSimonian–Laird method (DerSimonian and Laird 1986) be used to estimate τ^2. This method, historically, is one of the most popular estimation methods because it does not make any assumptions about the distribution of random effects and does not require iteration. But it may underestimate the true between-study variance, especially when the variability is large and the number of studies is small. This method may produce a negative value of τ^2 and is thus truncated at zero in that case.

  sjonkman specifies that the Sidik–Jonkman method (Sidik and Jonkman 2005) be used to estimate τ^2. This method always produces a nonnegative estimate of the between-study variance and thus does not need truncating at 0, unlike the other noniterative methods. Method sjonkman does not require iteration.

  hedges specifies that the Hedges method (Hedges 1983) be used to estimate τ^2. When the sampling variances of effect-size estimates can be estimated without bias, this estimator is exactly unbiased (before truncation), but it is not widely used in practice (Veroniki et al. 2016). Method hedges does not require iteration.

  hschmidt specifies that the Hunter–Schmidt method (Schmidt and Hunter 2015) be used to estimate τ^2. Although this estimator achieves a lower MSE than other methods, except ML, it is known to be negatively biased. Method hschmidt does not require iteration.

common specifies that a common-effect model be assumed for meta-analysis; see *Common-effect ("fixed-effect") model* in [META] **Intro**. It uses the inverse-variance estimation method; see *Meta-analysis estimation methods* in [META] **Intro**. Also see the discussion in [META] **meta data** about common-effect versus fixed-effects models.

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.

cefemethod is one of mhaenszel or invvariance (synonym ivariance). Below, we provide a short description for each method. Also see *Declaring a meta-analysis estimation method* in [META] **meta data**.

   mhaenszel is available only for a two-group comparison of binary outcomes. It specifies a meta-analysis using the Mantel–Haenszel method to estimate the overall effect size. This method is the default for effect sizes lnoratio, lnrratio, and rdiff but is not available for effect size lnorpeto.

   invvariance specifies a meta-analysis using the inverse-variance method to estimate the overall effect size. This method is available for all types of analyses and effect sizes. It is the default for a two-group comparison of continuous outcomes, for a two-group comparison of binary outcomes using effect size lnorpeto, and for estimating a single proportion (or prevalence).

   ivariance is a synonym for invvariance.

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.

   cefemethod is one of mhaenszel or invvariance (synonym ivariance); see descriptions above.

fixed specifies that a fixed-effects model be assumed for meta-analysis; see *Fixed-effects model* in [META] **Intro**. It uses the inverse-variance estimation method; see *Meta-analysis estimation methods* in [META] **Intro**. Also see the discussion in [META] **meta data** about fixed-effects versus common-effect models.

⎡ Options ⎤

zerocells(*zcspec*) is for use with binary outcomes when the effect size is either lnoratio or lnrratio for the two-sample case or either logitprop or proportion for the one-sample case. It specifies the adjustment to be used for the cells in the presence of zero cells. The cells are the values of variables *n11*, *n12*, *n21*, and *n22* for the two-sample case and the number of successes and the number of failures for the one-sample case. The adjustment is applied during computation—the original data are not modified. The default is zerocells(0.5, only0); it adds 0.5 to all cells of studies with at least one zero cell. To request no adjustment, specify zerocells(none). More generally, the syntax of *zcspec* is

   # [, *zcadj*]

where *#* is the adjustment value, also known as the continuity-correction value in the meta-analysis literature, and *zcadj* is only0 or allif0.

   only0 specifies that *#* be added to all cells of only those studies with at least one zero cell. For the two-sample case, during computation, *#* is added to each observation defined by variables *n11*, *n12*, *n21*, and *n22* if that observation contains a value of zero in any of those variables. For the one-sample case, *#* is added to all values (cells) corresponding to zero successes and to zero failures.

   allif0 specifies that *#* be added to all cells of all studies but only if there is at least one study with a zero cell. For the two-sample case, during computation, *#* is added to all values of variables *n11*, *n12*, *n21*, and *n22* but only if there is a zero value in one of the four variables. For the one-sample case, *#* is added to all cells (number of successes and number of failures) if at least one study contains zero successes or zero failures.

For the effect size lnoratio, *zcspec* may also be tacc, which implements the treatment-arm continuity correction of Sweeting, Sutton, and Lambert (2004). This method estimates the group-specific adjustment values from the data to minimize the bias of the overall odds-ratio estimator in the presence of zero cells. This method is recommended when the groups are unbalanced.

studylabel(*varname*) specifies a string variable containing labels for the individual studies to be used in all applicable meta-analysis output. The default study labels are Study 1, Study 2, . . . , Study K, where K is the total number of studies in the meta-analysis.

eslabel(*string*) specifies that *string* be used as the effect-size label in all relevant meta-analysis output. The default label is Effect size.

level(#) specifies the confidence level, as a percentage, for confidence intervals. It will be used by all subsequent meta-analysis commands when computing confidence intervals. The default is level(95) or as set by set level; see [R] level. After the declaration, you can specify level() with meta update to update the confidence level to be used throughout the rest of the meta-analysis session. You can also specify level() directly with the meta commands to modify the confidence level, temporarily, during the execution of the command.

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

Remarks and examples

stata.com

Remarks are presented under the following headings:

> *Meta-analysis for two-group comparison of binary outcomes*
> *Meta-analysis for two-group comparison of continuous outcomes*
> *Meta-analysis for estimating a single proportion*

meta esize computes various effect sizes, their standard errors, and CIs for continuous and binary outcomes from the summary data provided for each study. It then declares the computed effect-size data as the meta data; see [META] **meta data**. Different types of effect sizes may be specified in the esize() option. They depend on the type of analysis and outcome, so we describe them separately for various situations below, together with other data-specific options. Also see *Declaring meta-analysis information* in [META] **meta data**.

Meta-analysis for two-group comparison of binary outcomes

Meta-analysis is often used with studies comparing two groups. The first group is commonly referred to as the experimental or treatment group. The second group is commonly referred to as the control group.

For two-sample binary data, each study typically reports cell counts from the following 2×2 table.

group	success	failure	size
treatment	n_{11}	n_{12}	$n_1 = n_{11} + n_{12}$
control	n_{21}	n_{22}	$n_2 = n_{21} + n_{22}$

The cells of the table are composed of the numbers of "successes" and "failures" within each of the comparison groups. If a subject experiences an event of interest, it is a success; otherwise, it is a failure. Thus, the summary data for a two-group comparison of binary outcomes include the above 2×2 table for each study.

In this case, `meta esize` requires that four variables be specified containing the numbers of successes and failures in the treatment and control groups.

The goal of each study is to compare the probabilities of a success between the two groups. Various effect-size measures can be used for the comparison. For two-sample binary data, `meta esize` provides the following effect sizes: log odds-ratios (including Peto's log odds-ratios), the default; log risk-ratios; and risk differences. These are specified, respectively, as `lnoratio`, `lnorpeto`, `lnrratio`, and `rdiff` in the `esize()` option.

As described in *Declaring a meta-analysis model* in [META] **meta data**, you can choose between a random-effects, a fixed-effects, or a common-effect model. You can also choose from a number of estimation methods that are specific to the chosen model. For fixed-effects and common-effect models, in addition to the inverse-variance method, the Mantel–Haenszel method is available (and is the default) with effect sizes `lnoratio`, `lnrratio`, and `rdiff`; see *Declaring a meta-analysis estimation method* in [META] **meta data** and *Meta-analysis estimation methods* in [META] **Intro** for details.

Zero cell counts are known to create computational difficulties for odds ratios and risk ratios. A common solution is to add a small number, say, 0.5, to all cells of tables containing zero cells. This and other zero-cells adjustments are available in the `zerocells()` option.

Let's now look at several examples. Consider the following fictional meta-analysis dataset:

```
. use https://www.stata-press.com/data/r18/metaesbin
(Fictional data for binary outcomes)

. describe

Contains data from https://www.stata-press.com/data/r18/metaesbin.dta
 Observations:              4                      Fictional data for binary
                                                     outcomes
    Variables:              5                      23 Apr 2022 12:14
```

Variable name	Storage type	Display format	Value label	Variable label
study	str7	%9s		Study label
tdead	byte	%9.0g		Deaths in treatment group
tsurv	int	%9.0g		Survivors in treatment group
cdead	byte	%9.0g		Deaths in control group
csurv	int	%9.0g		Survivors in control group

```
Sorted by:
```

We will use this dataset to demonstrate how to compute effect sizes, specify different meta-analysis models, and adjust for zero cells with two-sample binary data.

▷ Example 1: A simple case

When working with meta-analysis data that do not have precomputed effect sizes, we can choose to compute effect sizes in a few different ways such as odds ratios and risk ratios. Using the simplest syntactical specification, we can compute the effect sizes, their standard errors, and the corresponding confidence intervals by specifying the number of successes and failures for one group, as well as the successes and failures for the second group, in that order.

```
. meta esize tdead tsurv cdead csurv

Meta-analysis setting information

 Study information
    No. of studies: 4
       Study label: Generic
        Study size: _meta_studysize
      Summary data: tdead tsurv cdead csurv

      Effect size
              Type: lnoratio
             Label: Log odds-ratio
          Variable: _meta_es
   Zero-cells adj.: 0.5, only0

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

  Model and method
             Model: Random effects
            Method: REML
```

The output indicates that there are 4 studies in the meta-analysis and, by default, a random-effects meta-analysis is to be assumed, where the heterogeneity parameter τ^2 is estimated via the REML method. The default computed effect size is the log odds-ratio. meta esize creates multiple system variables (see *System variables* in [META] **meta data**) that store the effect-size values, their standard errors, and the upper and lower limits of the CIs for the effect sizes.

We can now use, for example, meta summarize to list the individual log odds-ratios and the overall log odds-ratio, which is denoted as theta.

```
. meta summarize

  Effect-size label: Log odds-ratio
        Effect size: _meta_es
          Std. err.: _meta_se

Meta-analysis summary                    Number of studies =     4
Random-effects model                     Heterogeneity:
Method: REML                                        tau2 =  1.4417
                                                  I2 (%) =   69.33
                                                      H2 =    3.26
```

Study	Log odds-ratio	[95% conf. interval]		% weight
Study 1	-0.600	-2.079	0.879	27.80
Study 2	0.351	-2.510	3.212	15.65
Study 3	0.778	-0.031	1.586	34.69
Study 4	-2.567	-4.638	-0.495	21.85
theta	-0.403	-1.869	1.063	

```
Test of theta = 0: z = -0.54                    Prob > |z| = 0.5899
Test of homogeneity: Q = chi2(3) = 9.93            Prob > Q = 0.0192
```

See [META] **meta summarize** for details.

If we have a variable that stores the labels for each study, perhaps noting the study authors or journal, we can specify it in the studylabel() option with meta esize. Because we do not have such a variable in this dataset, each study is denoted generically by Study #. See example 4 in [META] **meta set** for an example of how to specify the study label and effect-size label.

◁

▷ Example 2: Specify the effect size

The default is to compute the log odds-ratio for the effect size. To specify another metric, we can use the esize() option. For example, below we use the risk ratio (on the log scale) as our effect size by specifying esize(lnrratio):

```
. meta esize tdead tsurv cdead csurv, esize(lnrratio)
Meta-analysis setting information
  Study information
     No. of studies: 4
        Study label: Generic
         Study size: _meta_studysize
       Summary data: tdead tsurv cdead csurv
         Effect size
                Type: lnrratio
               Label: Log risk-ratio
            Variable: _meta_es
    Zero-cells adj.: 0.5, only0
           Precision
           Std. err.: _meta_se
                  CI: [_meta_cil, _meta_ciu]
            CI level: 95%
   Model and method
               Model: Random effects
              Method: REML
```

◁

▷ Example 3: Sparse data and adjustments for zero cells

Note that when we list the data, one of the studies has zero deaths.

```
. list tdead tsurv cdead csurv
```

	tdead	tsurv	cdead	csurv
1.	2	116	17	541
2.	0	15	15	682
3.	8	61	37	614
4.	1	421	9	291

By default, `meta esize` adds a constant value of 0.5 (that is, option `zerocells(0.5, only0)` is assumed) to each cell of a study that has a zero cell; see `Zero-cells adj.:` in the output of `meta set` in example 1. We can modify this adjustment by specifying a different constant factor. For example, we might add 0.003 to each zero cell:

```
. meta esize tdead tsurv cdead csurv, zerocells(.003)
Meta-analysis setting information

 Study information
    No. of studies: 4
       Study label: Generic
        Study size: _meta_studysize
      Summary data: tdead tsurv cdead csurv

       Effect size
              Type: lnoratio
             Label: Log odds-ratio
          Variable: _meta_es
   Zero-cells adj.: .003, only0

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

   Model and method
             Model: Random effects
            Method: REML
```

Or we may instead choose a different type of continuity correction, for example, the treatment-arm continuity correction (TACC), which we specify as `zerocells(tacc)`:

```
. meta esize tdead tsurv cdead csurv, zerocells(tacc)
Meta-analysis setting information

 Study information
    No. of studies: 4
       Study label: Generic
        Study size: _meta_studysize
      Summary data: tdead tsurv cdead csurv

       Effect size
              Type: lnoratio
             Label: Log odds-ratio
          Variable: _meta_es
   Zero-cells adj.: tacc

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

   Model and method
             Model: Random effects
            Method: REML
```

Note that this option can be specified only when using the log odds-ratio as the effect size.

◁

▷ Example 4: Specify the meta-analysis model

In the examples above, we have been using the default random-effects model, but we could specify a different model. For example, we can use a common-effect model using the Mantel–Haenszel method to estimate the overall effect size:

```
. meta esize tdead tsurv cdead csurv, common(mhaenszel)
Meta-analysis setting information
 Study information
    No. of studies: 4
       Study label: Generic
        Study size: _meta_studysize
      Summary data: tdead tsurv cdead csurv
       Effect size
              Type: lnoratio
             Label: Log odds-ratio
          Variable: _meta_es
   Zero-cells adj.: 0.5, only0
          Precision
          Std. err.: _meta_se
                CI: [_meta_cil, _meta_ciu]
          CI level: 95%
  Model and method
             Model: Common effect
            Method: Mantel–Haenszel
```

In the above, we could have specified simply `common` because the Mantel–Haenszel method is the default for a common-effect model with log odds-ratios.

◁

Meta-analysis for two-group comparison of continuous outcomes

We can also use meta-analysis to compare two groups for continuous outcomes. As before, the first group is commonly referred to as the experimental or treatment group, and the second group is commonly referred to as the control group.

For a two-group comparison of continuous outcomes, each study often reports the numbers of observations, means, and standard deviations in the two groups. Various effect sizes are then computed from these summary data for each study. Thus, to compute effect sizes for two-sample continuous data, `meta esize` requires that six variables be specified containing the numbers of observations, means, and standard deviations of the treatment and control groups. The supported effect sizes are the raw mean difference, `esize(mdiff)`, and standardized mean differences: Hedges's g, `esize(hedgesg)` (the default); Cohen's d, `esize(cohend)`; and Glass's Δs, `esize(glassdelta2)` and `esize(glassdelta1)`; see *Methods and formulas* for their definitions.

As described in *Declaring a meta-analysis model* in [META] **meta data**, you can choose between a random-effects, a fixed-effects, or a common-effect model. You can also choose from several estimation methods for random-effects models. Fixed-effects and common-effect models assume the inverse-variance estimation method. Also see *Declaring a meta-analysis estimation method* in [META] **meta data** and *Meta-analysis estimation methods* in [META] **Intro** for details.

Let's now demonstrate several usages of `meta esize` for a two-group comparison of continuous outcomes. Consider the following fictional meta-analysis dataset:

```
. use https://www.stata-press.com/data/r18/metaescnt, clear
(Fictional summary data for continuous outcomes)

. describe

Contains data from https://www.stata-press.com/data/r18/metaescnt.dta
 Observations:              10                   Fictional summary data for
                                                   continuous outcomes
     Variables:              6                   19 Apr 2022 14:00
```

Variable name	Storage type	Display format	Value label	Variable label
n1	byte	%9.0g		Study sizes of group 1
m1	float	%9.0g		Means of group 1
sd1	float	%9.0g		Std. dev. of group 1
n2	byte	%9.0g		Study sizes of group 2
m2	float	%9.0g		Means of group 2
sd2	float	%9.0g		Std. dev. of group 2

```
Sorted by:
```

We will use this dataset to demonstrate different usages of the `meta esize` command with continuous-outcomes meta-analysis data.

▷ Example 5: The assumed model

In the simplest specification, `meta esize` requires that we specify the sample sizes, means, and standard deviations for each group in the meta-analysis.

```
. meta esize n1 m1 sd1 n2 m2 sd2
Meta-analysis setting information
 Study information
    No. of studies: 10
        Study label: Generic
         Study size: _meta_studysize
       Summary data: n1 m1 sd1 n2 m2 sd2

        Effect size
               Type: hedgesg
              Label: Hedges's g
           Variable: _meta_es
   Bias correction: Approximate

          Precision
          Std. err.: _meta_se
    Std. err. adj.: None
                 CI: [_meta_cil, _meta_ciu]
           CI level: 95%

  Model and method
              Model: Random effects
             Method: REML
```

We see from the output that the Hedges's g standardized mean difference is used for the effect size, and, as for binary outcomes, a random-effects REML model is assumed. See *Meta settings with meta esize* in [META] **meta data** for a detailed description of all settings for this dataset.

◁

▷ Example 6: Selecting an effect size

If we do not feel the need to standardize the mean differences, we could instead use the raw mean difference as the effect size by specifying esize(mdiff).

```
. meta esize n1 m1 sd1 n2 m2 sd2, esize(mdiff)
Meta-analysis setting information
 Study information
    No. of studies: 10
        Study label: Generic
         Study size: _meta_studysize
       Summary data: n1 m1 sd1 n2 m2 sd2

        Effect size
               Type: mdiff
              Label: Mean diff.
           Variable: _meta_es

           Precision
           Std. err.: _meta_se
    Std. err. adj.: None
                 CI: [_meta_cil, _meta_ciu]
           CI level: 95%

  Model and method
              Model: Random effects
             Method: REML
```

◁

▷ Example 7: Specifying different meta-analysis models and methods

Rather than using the default REML estimation method, we may want to use a different method, such as the DerSimonian–Laird method. We can specify this method in the random() option.

```
. meta esize n1 m1 sd1 n2 m2 sd2, random(dlaird)
Meta-analysis setting information
 Study information
    No. of studies: 10
        Study label: Generic
         Study size: _meta_studysize
       Summary data: n1 m1 sd1 n2 m2 sd2

        Effect size
               Type: hedgesg
              Label: Hedges's g
           Variable: _meta_es
    Bias correction: Approximate

           Precision
           Std. err.: _meta_se
    Std. err. adj.: None
                 CI: [_meta_cil, _meta_ciu]
           CI level: 95%

  Model and method
              Model: Random effects
             Method: DerSimonian–Laird
```

Or, instead of the random-effects model, we may specify a fixed-effects model, which implies the inverse-variance estimation method.

```
. meta esize n1 m1 sd1 n2 m2 sd2, fixed
Meta-analysis setting information
 Study information
    No. of studies: 10
        Study label: Generic
         Study size: _meta_studysize
       Summary data: n1 m1 sd1 n2 m2 sd2

         Effect size
                Type: hedgesg
               Label: Hedges's g
            Variable: _meta_es
   Bias correction: Approximate

           Precision
           Std. err.: _meta_se
     Std. err. adj.: None
                  CI: [_meta_cil, _meta_ciu]
            CI level: 95%

  Model and method
               Model: Fixed effects
              Method: Inverse-variance
```

◁

Meta-analysis for estimating a single proportion

Meta-analysis is also used to estimate an overall proportion (or prevalence) from one-sample binary data by pooling proportions from single-arm studies whenever this is sensible.

The data contain the number of successes (or the number of events) and the study sample size for each study. (Success is a generic term and occurs when a subject experiences an event of interest.) To estimate a proportion, `meta esize` provides the following effect sizes: Freeman–Tukey-transformed proportions (the default), logit-transformed proportions, and untransformed (raw) proportions. These are specified, respectively, as `ftukeyprop`, `logitprop`, and `proportion` in the `esize()` option.

As described in *Declaring a meta-analysis model* in [META] **meta data**, you can choose between a random-effects, a fixed-effects, or a common-effect model. You can also choose from several estimation methods for random-effects models. Fixed-effects and common-effect models assume the inverse-variance estimation method. Also see *Declaring a meta-analysis estimation method* in [META] **meta data** and *Meta-analysis estimation methods* in [META] **Intro** for details.

Let's now look at several examples. Consider the following fictional meta-analysis dataset:

```
. use https://www.stata-press.com/data/r18/metaesprop, clear
(Fictional summary data to estimate proportion)
. describe
Contains data from https://www.stata-press.com/data/r18/metaesprop.dta
 Observations:              6                    Fictional summary data to
                                                   estimate proportion
     Variables:             3                    26 Apr 2022 11:14

Variable       Storage   Display    Value
    name          type    format    label      Variable label

study           str7     %9s                   Study label
nsucc           byte     %9.0g                 Number of successes
ssize           int      %9.0g                 Study sample size

Sorted by:
```

We will use this dataset to demonstrate different usages of the `meta esize` command to declare the data for meta-analysis of a single proportion.

▷ Example 8: The default setting

In its most basic form, `meta esize` requires that we specify the number of successes (`nsucc`) and the study sample sizes (`ssize`).

```
. meta esize nsucc ssize
Meta-analysis setting information
 Study information
    No. of studies: 6
       Study label: Generic
        Study size: _meta_studysize
      Summary data: nsucc ssize

      Effect size
              Type: ftukeyprop
             Label: Freeman-Tukey's p
          Variable: _meta_es

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

  Model and method
             Model: Random effects
            Method: REML
```

The output shows that the summary data are defined by variables `nsucc` and `ssize` and that, by default, the Freeman–Tukey-transformed proportion is used as the effect size. A random-effects REML model is assumed. Other settings are exactly as described in example 1.

◁

▷ Example 9: Specify the effect size

Instead of using the default Freeman–Tukey-transformed proportion, we can choose a different effect size, such as the logit-transformed proportion, using the `esize()` option.

```
. meta esize nsucc ssize, esize(logitprop)
Meta-analysis setting information
 Study information
    No. of studies: 6
       Study label: Generic
        Study size: _meta_studysize
      Summary data: nsucc ssize

      Effect size
              Type: logitprop
             Label: Logit proportion
          Variable: _meta_es
   Zero-cells adj.: None; no zero cells

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

  Model and method
             Model: Random effects
            Method: REML
```

The output differs from that in example 8 in the Effect size section. It now reflects that logit-transformed proportion is the effect size of choice instead of the default Freeman–Tukey-transformed proportion. There is also a new row for the zero-cells adjustment. This row did not show up in the output of example 8 because the Freeman–Tukey-transformed proportion does not need continuity correction. In our dataset, there are no zero cells, so the output in that row shows that no zero-cells adjustment was applied.

The logit-transformed proportion (and the untransformed proportion, esize(proportion)) should be avoided when there are study proportions that are close to 0 or 1.

◁

▷ Example 10: Specify an alternative meta-analysis model and method

Instead of using the default REML estimation method, you may specify an alternative random-effects method, such as the DL method. This can be done via the random() option.

```
. meta esize nsucc ssize, random(dlaird)
Meta-analysis setting information
 Study information
    No. of studies: 6
      Study label: Generic
       Study size: _meta_studysize
     Summary data: nsucc ssize
      Effect size
             Type: ftukeyprop
            Label: Freeman-Tukey's p
         Variable: _meta_es
         Precision
        Std. err.: _meta_se
               CI: [_meta_cil, _meta_ciu]
         CI level: 95%
   Model and method
            Model: Random effects
           Method: DerSimonian-Laird
```

Or perhaps you believe that your proportions are similar across the studies and that a common-effect model is adequate to synthesize the overall proportion. You may request a common-effect model with the inverse-variance method by specifying the common option.

```
. meta esize nsucc ssize, common
Meta-analysis setting information
 Study information
    No. of studies: 6
      Study label: Generic
       Study size: _meta_studysize
     Summary data: nsucc ssize
      Effect size
             Type: ftukeyprop
            Label: Freeman-Tukey's p
         Variable: _meta_es
         Precision
        Std. err.: _meta_se
               CI: [_meta_cil, _meta_ciu]
         CI level: 95%
   Model and method
            Model: Common effect
           Method: Inverse-variance
```

◁

Stored results

meta esize stores the following characteristics and system variables:

Characteristics

_dta[_meta_marker]	"_meta_ds_1"
_dta[_meta_K]	number of studies in the meta-analysis
_dta[_meta_studylabel]	name of string variable containing study labels or Generic
_dta[_meta_estype]	type of effect size; varies
_dta[_meta_eslabelopt]	eslabel(*eslab*), if specified
_dta[_meta_eslabel]	effect-size label from eslabel(); default varies
_dta[_meta_eslabeldb]	effect-size label for dialog box
_dta[_meta_esvardb]	_meta_es
_dta[_meta_level]	default confidence level for meta-analysis
_dta[_meta_esizeopt]	esize(*estype*), if specified
_dta[_meta_esopt_exact]	exact, if esize(, exact) is specified
_dta[_meta_esopt_holkinse]	holkinse, if esize(, holkinse) is specified
_dta[_meta_esopt_unequal]	unequal, if esize(, unequal) is specified
_dta[_meta_modellabel]	meta-analysis model label: Random effects, Common effect, or Fixed effects
_dta[_meta_model]	meta-analysis model: random, common, or fixed
_dta[_meta_methodlabel]	meta-analysis method label; varies by meta-analysis model
_dta[_meta_method]	meta-analysis method; varies by meta-analysis model
_dta[_meta_randomopt]	random(*remethod*), if specified
_dta[_meta_zcopt]	zerocells(*zcspec*), if specified
_dta[_meta_zcadj]	type of adjustment for zero cells, if zerocells() specified
_dta[_meta_zcvalue]	value added to cells to adjust for zero cells, if specified
_dta[_meta_show]	empty or nometashow
_dta[_meta_n1var]	name of group 1 sample-size variable; for two-sample continuous data
_dta[_meta_mean1var]	name of group 1 mean variable; for two-sample continuous data
_dta[_meta_sd1var]	name of group 1 std. dev. variable; for two-sample continuous data
_dta[_meta_n2var]	name of group 2 sample-size variable; for two-sample continuous data
_dta[_meta_mean2var]	name of group 2 mean variable; for two-sample continuous data
_dta[_meta_sd2var]	name of group 2 std. dev. variable; for two-sample continuous data
_dta[_meta_n11var]	name of *n11* variable; for two-sample binary data (contingency table)
_dta[_meta_n12var]	name of *n12* variable; for two-sample binary data (contingency table)
_dta[_meta_n21var]	name of *n21* variable; for two-sample binary data (contingency table)
_dta[_meta_n22var]	name of *n22* variable; for two-sample binary data (contingency table)
_dta[_meta_nsvar]	name of *ns* variable; for one-sample binary data (proportion)
_dta[_meta_nvar]	name of *n* variable; for one-sample binary data (proportion)
_dta[_meta_datatype]	data type; continuous, binary, or proportion
_dta[_meta_datavars]	variables specified with meta esize
_dta[_meta_setcmdline]	meta esize command line
_dta[_meta_ifexp]	*if* specification
_dta[_meta_inexp]	*in* specification

System variables

_meta_id	study ID variable
_meta_es	variable containing effect sizes
_meta_se	variable containing effect-size standard errors
_meta_cil	variable containing lower bounds of CIs for effect sizes
_meta_ciu	variable containing upper bounds of CIs for effect sizes
_meta_studylabel	string variable containing study labels
_meta_studysize	variable containing total sample size per study

Methods and formulas

Methods and formulas are presented under the following headings:

Effect sizes for two-group comparison of continuous outcomes

As we described in *Meta-analysis for two-group comparison of continuous outcomes*, meta-analysis often compares two groups: experimental (or treated) group and control group.

When the response (measurement) is continuous, studies typically report a mean and standard deviation for each group. For a given study, the following table denotes the underlying population parameters and the reported summary statistics (data) for each group.

	population		sample		
group	mean	sd	mean	sd	size
treatment	μ_1	σ_1	$\overline{x}_1$	s_1	n_1
control	μ_2	σ_2	$\overline{x}_2$	s_2	n_2

The majority of this section is based on Borenstein (2009).

Unstandardized mean difference

Consider the population mean difference

$$\theta = \mu_1 - \mu_2$$

For each study in the meta-analysis, meta esize with option esize(mdiff) estimates θ using the difference in sample means,

$$D = \overline{x}_1 - \overline{x}_2$$

The variance of D, assuming that the two population standard deviations are equal, is estimated by

$$\widehat{\text{Var}}(D) = \left(\frac{1}{n_1} + \frac{1}{n_2} \right) s^2$$

where s is the pooled sample standard deviation

$$s = \sqrt{\frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2}{n_1 + n_2 - 2}}$$

For unequal population standard deviations, use option `esize(mdiff, unequal)`; then the variance of D is estimated by

$$\widehat{\text{Var}}(D) = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$$

Unstandardized (raw) mean differences are not comparable across studies if the underlying means are measured on different scales.

Standardized mean difference

The standardized mean difference is

$$\theta = \frac{\mu_1 - \mu_2}{\sigma}$$

Note that θ does not depend on the scale of measurement. The definition of the standardized mean difference implicitly assumes that the population standard deviations, σ_1 and σ_2, are the same: $\sigma_1 = \sigma_2 = \sigma$.

`meta esize` with option `esize(cohend)` estimates θ using Cohen's d statistic (Cohen 1969, 1988),

$$d = \frac{\overline{x}_1 - \overline{x}_2}{s}$$

The estimated variance of d is given by

$$\widehat{\text{Var}}(d) = \frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2\,(n_1 + n_2)}$$

Hedges (1981) introduced an adjustment to Cohen's d for small samples that accounts for a small upward bias in the absolute value of θ. `meta esize` with option `esize(hedgesg, exact)` computes Hedges's g as

$$g = c(m) \times d$$

where $m = n_1 + n_2 - 2$ is the degrees of freedom used to estimate s and

$$c(m) = \frac{\Gamma\left(\frac{m}{2}\right)}{\sqrt{\frac{m}{2}}\Gamma\left(\frac{m-1}{2}\right)}$$

The adjustment $c(m)$ is less than 1 and approaches 1 as m gets large. The variance estimate of Hedges's g is

$$\widehat{\text{Var}}(g) = c(m)^2 \times \widehat{\text{Var}}(d)$$

Hedges (1981) also introduced an accurate approximation for $c(m)$ that has been traditionally used in meta-analysis. The approximation for $c(m)$ is

$$J = 1 - \frac{3}{4m - 1}$$

`meta esize` with option `esize(hedgesg)` computes Hedges's g using J for $c(m)$; thus,

$$g = J \times d$$

and

$$\widehat{\mathrm{Var}}(g) = J^2 \times \widehat{\mathrm{Var}}(d)$$

`meta esize` with option `esize(glassdelta2)` estimates θ using Glass's Δ (Smith and Glass 1977),

$$\Delta = \frac{\overline{x}_1 - \overline{x}_2}{s_2}$$

Notice that the standard deviation in the denominator is s_2, the sample standard deviation from the control group, which is considered to be a more reliable estimate of the common variance. The estimated variance of Δ is given by

$$\widehat{\mathrm{Var}}(\Delta) = \frac{n_1 + n_2}{n_1 n_2} + \frac{\Delta^2}{2\,(n_2 - 1)}$$

In the absence of the control group, such as in observational studies, Kline (2013), among others, suggests providing statistics standardized by the standard deviation of each group. Glass's Δ where standardization is based on the treatment group may be computed via option `esize(glassdelta1)`.

Alternative standard error estimators are available for Hedges's g and Cohen's d effect sizes.

Hedges and Olkin (1985, eq. 8, 80) provide another commonly used estimator for the variance of Hedges's g.

$$\widehat{\mathrm{Var}}(g) = \frac{n_1 + n_2}{n_1 n_2} + \frac{g^2}{2\,(m - 1.94)}$$

`meta esize` uses this formula when option `esize(hedgesg, holkinse)` is specified.

The alternative variance estimator of d is given by

$$\widehat{\mathrm{Var}}(d) = \frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2\,(n_1 + n_2 - 2)}$$

This variance estimator may be requested via option `esize(cohend, holkinse)`.

Effect sizes for two-group comparison of binary outcomes

As we described in *Meta-analysis for two-group comparison of binary outcomes*, meta-analysis often compares two groups: experimental (or treated) group and control group. When the response (measurement) is binary, each study typically reports cell counts from the following 2×2 table.

group	success	failure	size
treatment	a	b	$n_1 = a + b$
control	c	d	$n_2 = c + d$

Here, for simplicity, we use a different notation for the cell counts (a, b, c, and d) compared with the similar table in *Meta-analysis for two-group comparison of binary outcomes*.

For the treatment group, n_1 is assumed fixed, $a \sim \mathrm{binomial}(n_1, \pi_1)$, and π_1 is the probability of a success. For the control group, n_2 is assumed fixed, $c \sim \mathrm{binomial}(n_2, \pi_2)$, and π_2 is the probability of a success. The goal of each study is to compare the two success probabilities, π_1 and π_2.

Estimates of the success probabilities are $\widehat{\pi}_1 = a/n_1$ for the treatment group and $\widehat{\pi}_2 = c/n_2$ for the control group.

Odds ratio

`meta esize` with option `esize(lnoratio)` computes estimates of the log odds-ratios. Odds ratio is the ratio of the odds of a success in the treatment group over the odds of a success in the control group.

$$\text{OR} = \frac{\pi_1 / (1 - \pi_1)}{\pi_2 / (1 - \pi_2)}$$

The odds ratio is estimated by

$$\widehat{\text{OR}} = \frac{ad}{bc}$$

The distribution of $\widehat{\text{OR}}$ is typically skewed, but the natural logarithm of $\widehat{\text{OR}}$, $\ln(\widehat{\text{OR}})$, is asymptotically normally distributed. The estimate of the variance of $\ln(\widehat{\text{OR}})$ is

$$\widehat{\text{Var}}\left\{\ln(\widehat{\text{OR}})\right\} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$$

`meta esize` with option `esize(lnorpeto)` computes estimates of effect size using Peto's log odds-ratio (Peto et al. 1977; Yusuf et al. 1985). Peto's odds ratio and log odds-ratio are

$$\widehat{\text{OR}}^{\text{Peto}} = \exp\left\{\frac{a - E(a)}{\text{Var}(a)}\right\}$$

$$\ln\left(\widehat{\text{OR}}^{\text{Peto}}\right) = \frac{a - E(a)}{\text{Var}(a)}$$

where the expectation and variance of a are estimated assuming a hypergeometric distribution:

$$E(a) = \frac{(a + c)n_1}{n}$$

$$\text{Var}(a) = \frac{n_1 n_2 (a + c)(b + d)}{n^2 (n - 1)}$$

The variance estimate of $\ln\left(\widehat{\text{OR}}^{\text{Peto}}\right)$ is

$$\widehat{\text{Var}}\left\{\ln\left(\widehat{\text{OR}}^{\text{Peto}}\right)\right\} = \frac{1}{\text{Var}(a)}$$

See, for instance, Fleiss 1993, Fleiss, Levin, and Paik 2003, and Bradburn et al. (2007) for a discussion of potential bias of Peto's odds ratio and its performance in sparse data.

Risk ratio (rate ratio)

`meta esize` with option `esize(lnrratio)` computes estimates of the log risk-ratios. The risk ratio (RR), also known as the rate ratio or relative risk in the health sciences, is

$$\text{RR} = \frac{\pi_1}{\pi_2}$$

RR is estimated by

$$\widehat{\text{RR}} = \frac{a/n_1}{c/n_2}$$

Similarly to odds ratios, $\widehat{\text{RR}}$ typically has a skewed distribution, but the natural logarithm of $\widehat{\text{RR}}$, $\ln\left(\widehat{\text{RR}}\right)$, is asymptotically normally distributed. The estimate of the variance of $\ln\left(\widehat{\text{RR}}\right)$ is

$$\widehat{\text{Var}}\left\{\ln\left(\widehat{\text{RR}}\right)\right\} = \frac{1}{a} + \frac{1}{c} - \frac{1}{a + b} - \frac{1}{c + d}$$

Risk difference

`meta esize` with option `esize(rdiff)` computes estimates of the risk differences. The risk difference is

$$\text{RD} = \pi_1 - \pi_2$$

and is estimated by

$$\widehat{\text{RD}} = \frac{a}{n_1} - \frac{c}{n_2}$$

$\widehat{\text{RD}}$ is asymptotically normally distributed and is thus typically used without a transformation in meta-analysis.

The estimated variance of $\widehat{\text{RD}}$ is

$$\widehat{\text{Var}}\left(\widehat{\text{RD}}\right) = \frac{ab}{n_1^3} + \frac{cd}{n_2^3}$$

Zero-cells adjustments for two-sample case

The variance estimates of $\ln(\widehat{\text{OR}})$ and $\ln\left(\widehat{\text{RR}}\right)$ are not defined if there are any empty (zero count) cells in a 2×2 table. In this case, it is customary to add a small value, often referred to as "continuity correction", to each cell prior to computing the log odds- or risk-ratio.

By default, `meta esize` adds 0.5 to each cell of 2×2 tables containing empty cells (Gart and Zweifel 1967 and Gart, Pettigrew, and Thomas 1985). Alternatively, you can add a different number or add a number to each cell of all 2×2 tables, as long as there is at least one 2×2 table with zero cells; see option `zerocells()`.

For odds ratios, Sweeting, Sutton, and Lambert (2004) proposed the treatment-arm continuity correction (TACC) method, which estimates the continuity-correction values from the data separately for each group; see `zerocells(tacc)`.

Effect sizes for estimating a single proportion

As we described in *Meta-analysis for estimating a single proportion*, meta-analysis may be used to aggregate proportions of a certain event of interest in single-group or single-arm studies. Each study typically reports the number of successes (number of events), e, and the study sample size, n. The number of successes e is assumed to follow a binomial(n, p) distribution, where p is the probability of success. For details, see Barendregt et al. (2013) and Nyaga, Arbyn, and Aerts (2014).

Raw (untransformed) proportion

`meta esize` with option `esize(proportion)` computes estimates of proportions for each study and uses them as effect sizes in the meta-analysis. The proportion is estimated by

$$\widehat{p} = \frac{e}{n}$$

When the proportion p is near 0.5 and when n is sufficiently large, the binomial distribution of e is well approximated by the normal distribution, and a meta-analysis may be performed in the natural (untransformed) metric.

The estimated variance of $\widehat{p}$ is

$$\widehat{\mathrm{Var}}\left(\widehat{p}\right) = \frac{\widehat{p}\left(1-\widehat{p}\right)}{n}$$

Because the expression of the variance depends on $\widehat{p}$, meta-analysis of this effect size tends to assign artificially large weights for studies with $\widehat{p}$ close to 0 or 1. In this case, the variance of $\widehat{p}$ is close to 0, and the study weights, which are the inverse variances, will be large. Also, study-specific CI limits may fall outside the range of $[0, 1]$ and, in practice, are truncated when this happens.

Freeman–Tukey-transformed proportion

By default (or with option `esize(ftukeyprop)`), `meta esize` computes the Freeman–Tukey-transformed proportions and uses them as effect sizes in the meta-analysis. The Freeman–Tukey transformation is also known as the Freeman–Tukey double-arcsine transformation in the literature. The Freeman–Tukey-transformed proportion is given by

$$\widehat{p}_{\mathrm{FT}} = \mathtt{asin}\left(\sqrt{\frac{e}{n+1}}\right) + \mathtt{asin}\left(\sqrt{\frac{e+1}{n+1}}\right) \tag{1}$$

with the corresponding estimated variance

$$\widehat{\mathrm{Var}}\left(\widehat{p}_{\mathrm{FT}}\right) = \frac{1}{n+0.5}$$

This is a variance-stabilizing transformation (variance does not depend on e) and is particularly preferable when $\widehat{p}$ is close to 0 or 1. This transformation also addresses the issue of assigning artificially small or large weights to studies in the meta-analysis when e is close to 0 or n. And it guarantees that the back-transformed CIs (see *Inverse Freeman–Tukey transformation* in *Methods and formulas* in [META] **meta summarize**) fall within the $[0, 1]$ range.

Logit-transformed proportion

`meta esize` with option `esize(logitprop)` computes logit-transformed proportions and uses them as effect sizes in the meta-analysis. The logit-transformed proportion is estimated by

$$\mathtt{logit}\left(\widehat{p}\right) = \ln\left(\frac{\widehat{p}}{1-\widehat{p}}\right)$$

with the corresponding estimated variance

$$\widehat{\mathrm{Var}}\left\{\mathtt{logit}\left(\widehat{p}\right)\right\} = \frac{1}{n\widehat{p}} + \frac{1}{n-n\widehat{p}}$$

This transformation allows aggregating the proportions in a metric that is closer to normality and guarantees that the back-transformed CI limits (computed using the `invlogit()` function) are between 0 and 1 (inclusive). Because the expression of the variance depends on $\widehat{p}$, meta-analysis of this effect size tends to assign artificially low weights for studies with $\widehat{p}$ close to 0 or 1. In this case, the variances for such studies are large, and the study weights, which are the inverse variances, will be low.

Zero-cells adjustments for one-sample case

When a study reports a zero cell (zero successes or zero failures), the variance of $\widehat{p}$ is equal to 0, and the variance of $\text{logit}(\widehat{p})$ is not defined. In this case, it is customary to add a small value, often referred to as "continuity correction", to each cell prior to computing the proportion or the logit-transformed proportion.

By default, `meta esize` adds 0.5 to each cell of studies containing zero cells (Gart and Zweifel 1967 and Gart, Pettigrew, and Thomas 1985). In other words, for a study reporting zero cells, the number of successes, e, will be incremented by 0.5, the number of failures will be incremented by 0.5, and therefore, the total sample size, n, will increase by 1.

Alternatively, you can add a different number or add a number to each cell of all studies, as long as there is at least one study with zero cells; see option `zerocells()`.

Confidence intervals for effect sizes

For the jth study in a given meta-analysis, let $\widehat{\theta}_j$ be one of the effect-size estimators described above; then the asymptotic $100(1 - \alpha)\%$ confidence interval computed by `meta esize` is

$$\widehat{\theta}_j \pm z_{1-\alpha/2}\sqrt{\widehat{\text{Var}(\widehat{\theta}_j)}}$$

where $z_{1-\alpha/2}$ is the usual critical value from the standard normal distribution.

References

Barendregt, J. J., S. A. Doi, Y. Y. Lee, R. E. Norman, and T. Vos. 2013. Meta-analysis of prevalence. *Journal of Epidemiology and Community Health* 67: 974–978. https://doi.org/10.1136/jech-2013-203104.

Berkey, C. S., D. C. Hoaglin, F. Mosteller, and G. A. Colditz. 1995. A random-effects regression model for meta-analysis. *Statistics in Medicine* 14: 395–411. https://doi.org/10.1002/sim.4780140406.

Borenstein, M. 2009. Effect sizes for continuous data. In *The Handbook of Research Synthesis and Meta-Analysis*, ed. H. Cooper, L. V. Hedges, and J. C. Valentine, 2nd ed., 221–235. New York: Russell Sage Foundation.

Bradburn, M. J., J. J. Deeks, J. A. Berlin, and A. R. Localio. 2007. Much ado about nothing: A comparison of the performance of meta-analytical methods with rare events. *Statistics in Medicine* 26: 53–77. https://doi.org/10.1002/sim.2528.

Cohen, J. 1969. *Statistical Power Analysis for the Behavioral Sciences.* New York: Academic Press.

———. 1988. *Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. Hillsdale, NJ: Erlbaum.

DerSimonian, R., and N. M. Laird. 1986. Meta-analysis in clinical trials. *Controlled Clinical Trials* 7: 177–188. https://doi.org/10.1016/0197-2456(86)90046-2.

Fleiss, J. L. 1993. Review papers: The statistical basis of meta-analysis. *Statistical Methods in Medical Research* 2: 121–145. https://doi.org/10.1177/096228029300200202.

Fleiss, J. L., B. Levin, and M. C. Paik. 2003. *Statistical Methods for Rates and Proportions.* 3rd ed. New York: Wiley.

Freeman, M. F., and J. W. Tukey. 1950. Transformations related to the angular and the square root. *Annals of Mathematical Statistics* 21: 607–611. https://doi.org/10.1214/aoms/1177729756.

Gart, J. J., H. M. Pettigrew, and D. G. Thomas. 1985. The effect of bias, variance estimation, skewness and kurtosis of the empirical logit on weighted least squares analyses. *Biometrika* 72: 179–190. https://doi.org/10.2307/2336348.

Gart, J. J., and J. R. Zweifel. 1967. On the bias of various estimators of the logit and its variance with application to quantal bioassay. *Biometrika* 54: 181–187. https://doi.org/10.2307/2333861.

Hardy, R. J., and S. G. Thompson. 1996. A likelihood approach to meta-analysis with random effects. *Statistics in Medicine* 15: 619–629. https://doi.org/10.1002/(SICI)1097-0258(19960330)15:6⟨619::AID-SIM188⟩3.0.CO;2-A.

Hedges, L. V. 1981. Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics* 6: 107–128. https://doi.org/10.2307/1164588.

———. 1983. A random effects model for effect sizes. *Psychological Bulletin* 93: 388–395. http://doi.org/10.1037/0033-2909.93.2.388.

Hedges, L. V., and I. Olkin. 1985. *Statistical Methods for Meta-Analysis*. Orlando, FL: Academic Press.

Kline, R. B. 2013. *Beyond Significance Testing: Statistics Reform in the Behavioral Sciences*. 2nd ed. Washington, DC: American Psychological Association.

Nyaga, V. N., M. Arbyn, and M. Aerts. 2014. Metaprop: A Stata command to perform meta-analysis of binomial data. *Archives of Public Health* 72. https://doi.org/10.1186/2049-3258-72-39.

Paule, R. C., and J. Mandel. 1982. Consensus values and weighting factors. *Journal of Research of the National Bureau of Standards* 87: 377–385. http://doi.org/10.6028/jres.087.022.

Peto, R., M. C. Pike, P. Armitage, N. E. Breslow, D. R. Cox, S. V. Howard, N. Mantel, K. McPherson, J. Peto, and P. G. Smith. 1977. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *British Journal of Cancer* 35: 1–39. https://doi.org/10.1038/bjc.1977.1.

Raudenbush, S. W. 2009. Analyzing effect sizes: Random-effects models. In *The Handbook of Research Synthesis and Meta-Analysis*, ed. H. Cooper, L. V. Hedges, and J. C. Valentine, 2nd ed., 295–316. New York: Russell Sage Foundation.

Schmidt, F. L., and J. E. Hunter. 2015. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. 3rd ed. Thousand Oaks, CA: Sage.

Sidik, K., and J. N. Jonkman. 2005. A note on variance estimation in random effects meta-regression. *Journal of Biopharmaceutical Statistics* 15: 823–838. https://doi.org/10.1081/BIP-200067915.

Smith, M. L., and G. V. Glass. 1977. Meta-analysis of psychotherapy outcome studies. *American Psychologist* 32: 752–760. http://doi.org/10.1037/0003-066X.32.9.752.

Sweeting, M. J., A. J. Sutton, and P. C. Lambert. 2004. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Statistics in Medicine* 23: 1351–1375. https://doi.org/10.1002/sim.1761.

Veroniki, A. A., D. Jackson, W. Viechtbauer, R. Bender, J. Bowden, G. Knapp, O. Kuss, J. P. T. Higgins, D. Langan, and G. Salanti. 2016. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Research Synthesis Methods* 7: 55–79. https://doi.org/10.1002/jrsm.1164.

Yusuf, S., R. Peto, J. Lewis, R. Collins, and P. Sleight. 1985. Beta blockade during and after myocardial infarction: An overview of the randomized trials. *Progress in Cardiovascular Diseases* 27: 335–371. https://doi.org/10.1016/S0033-0620(85)80003-7.

Also see