THE STATA JOURNAL

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*), *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Two-stage individual participant data meta-analysis and generalized forest plots

David J. Fisher
MRC Clinical Trials Unit at University College London
London, UK
d.fisher@ucl.ac.uk

Abstract. In this article, I describe a command, `ipdmetan`, that facilitates two-stage individual participant data meta-analysis of any measure of effect and its standard error by fitting a specified model to data from each study. The command can estimate random effects and heterogeneity statistics and include additional covariates and interactions. If individual participant data are available for certain studies and aggregate data for others, `ipdmetan` allows them to be combined in one analysis. This command can produce detailed and flexible forest plots, including ones outside the context of formal meta-analysis.

Keywords: st0384, ipdmetan, forestplot, admetan, ipdover, meta-analysis, forest plot, individual participant data, survival, two-stage, time-to-event, prefix command, subgroup

1 Introduction

Meta-analysis is a statistical technique for combining results from multiple independent studies, usually to estimate one overall effect with increased precision (Deeks, Altman, and Bradburn 2001). Stata has various tools for performing meta-analysis, many of which are described in a previous collection of articles (Sterne 2009). These have focused on the pooling of study results in the form of aggregate (published summary) data (AD)—traditionally the most common form of meta-analysis. However, individual participant (or patient) data (IPD) are increasingly used and permit a greater range of possible analyses while minimizing bias from heterogeneity of analysis protocols between studies (for example, patient exclusions or the precise model used) and maximizing data completeness and follow-up (Stewart and Tierney 2002).

Both "one-stage" and "two-stage" approaches exist for analyzing IPD; both are commonly used and are consistent estimators of the pooled effect. Although a one-stage estimator is always at least as efficient as its equivalent two-stage estimator for linear models where the full covariance matrix is directly observed (Mathew and Nordström 2010), it remains unclear whether this holds for nonlinear outcomes (for example, time-to-event) and situations where the covariate matrix is only an estimate. Moreover, it is unclear whether such differences are great enough to be statistically or clinically relevant. For instance, simulation studies (Tudur Smith and Williamson 2007; Bowden et al. 2011) and case studies (Stewart et al. 2012) have shown little or no differences between one-stage and two-stage models under specific conditions. Furthermore, for situations such

as random-effects analysis of time-to-event data, one-stage solutions are just becoming available (Bowden et al. 2011; Crowther, Look, and Riley 2014). Partly for this reason, the most common approach used with time-to-event data is the two-stage approach (Simmonds et al. 2005).

In practice, the two-stage approach is appealing in its simplicity and because it can call on extensive AD meta-analysis literature, including the *Cochrane Handbook* (Higgins and Green 2011). Two-stage models can offer quick and unbiased answers, especially for reviews of randomized controlled trials (RCTs), where studies are wholly independent and confounding is minimized. Also, for some situations, a one-stage model would require innovative techniques. Furthermore, studies without IPD available may be included as a sensitivity analysis (Riley et al. 2008) by simply including AD from publications and by using secondary requests for additional data (for example, an effect estimate adjusted for extra covariates) after an initial request for IPD has failed. Nevertheless, two-stage IPD analysis in Stata has not previously been straightforward. Now, `ipdmetan` provides a comprehensive two-stage package.

One motivation for collecting IPD is the possibility of assessing differential responses to treatment—in other words, of analyzing treatment-covariate interactions. In previous work (Fisher et al. 2011), we discussed the fundamentally different approaches to the analysis of such data and concluded that a two-stage approach is again suitable in many situations, especially compared with older methods. In future work, we hope to expand on those suggestions and discuss how trialists might facilitate meta-analysis of interactions, even if they are unwilling or unable to provide IPD, by presenting subgroup analyses clearly and consistently and thus allowing aggregate estimates to be calculated.

The standard method of presenting meta-analysis results is a forest plot, in which the results from each study are plotted together with the overall result. Previous Stata AD meta-analysis commands produce forest plots as standard, and such plots transfer naturally to two-stage IPD analysis. Results from one-stage models may also be plotted in this way, but with greater complications. The `ipdforest` command (Kontopantelis and Reeves 2013) is one approach to this; another might be to present two-stage study results alongside a one-stage pooled result. Forest plots for `ipdmetan` are created using code based heavily on that written for the `metan` command (Harris et al. 2008) but modified and updated for increased flexibility. They are produced by the program `forestplot`, which is called by `ipdmetan` but can also be run by itself, giving the user more control.

Finally, the `ipdmetan` package also includes the following "wrapper" programs, which I do not describe in detail here: `admetan`, an AD routine similar to `metan`; `petometan`, a routine for performing IPD meta-analysis of time-to-event data using the Peto logrank approach (Yusuf et al. 1985); and `ipdover`, a routine that creates forest plots of subgroups within one trial.

2 Two-stage IPD meta-analysis

2.1 Basic principles

Two-stage meta-analysis methodology involves first fitting the desired model to the data from each study in turn. This gives study-effect estimates and variances, which are denoted here by θ_i and σ_i^2, respectively. These are assumed to be observed without error in the second (pooling) stage. Let θ denote the true pooled effect to be estimated by $\widehat{\theta}$. A suitable reference for most of what follows is Deeks, Altman, and Bradburn (2001), with extra references given where appropriate.

ipdmetan uses inverse-variance weighting, with weights denoted by $w_i = 1/\sigma_i^2$. The Peto (or log-rank) method for time-to-event outcomes is equivalent to an inverse-variance analysis, and it can be performed using ipdmetan as described in section 4.5. The built-in commands cc or tabodds can be used to perform IPD meta-analysis with Mantel–Haenszel weighting for binary outcomes.

The fixed-effects inverse-variance weighted pooled effect and its variance are calculated as

$$\widehat{\theta} = \frac{\sum\limits_i w_i \theta_i}{\sum\limits_i w_i}, \; \mathrm{Var}\left(\widehat{\theta}\right) = \frac{1}{\sum\limits_i w_i}$$

The standard Cochran heterogeneity statistic Q is defined as

$$Q = \sum_i w_i \left(\theta_i - \widehat{\theta}\right)^2$$

and is distributed as chi-squared with $k - 1$ degrees of freedom (d.f.), where k is the number of included studies.

Given an estimate of between-study variance τ^2, the random-effects inverse-variance weighted pooled effect (and variance) is calculated by replacing the fixed-effects weights w_i with random-effects weights $w_i^* = 1/(\sigma_i^2 + \tau^2)$, which incorporate the between-study variance τ^2. The most common estimator of τ^2 is the DerSimonian–Laird (DL) estimator τ_{DL}^2 (see section 2.2).

In addition to Q and τ^2, two further heterogeneity measures are commonly used: I^2 and H_M^2 (introduced by Higgins and Thompson [2002] and Mittlböck and Heinzl [2006]). We first define the "typical" or "average" within-study variance (Higgins and Thompson 2002) as

$$s^2 = \frac{\sum\limits_i w_i(k-1)}{\left(\sum\limits_i w_i\right)^2 - \sum\limits_i w_i^2}$$

I^2 and H_M^2 can then be defined as follows:

$$I^2 = \frac{\tau^2}{\tau^2 + s^2}; \; H_M^2 = \frac{\tau^2}{s^2}$$

If $\tau^2 = \tau^2_{\mathrm{DL}}$, then the algebra may be rearranged to form the alternative formula

$$I^2 = \frac{Q - \mathrm{d.f.}}{Q}; \; H^2_M = \frac{Q - \mathrm{d.f.}}{\mathrm{d.f.}}$$

where $\mathrm{d.f.} = k - 1$ is the number of d.f. associated with Q (as defined in section 2.1).

Therefore, I^2 and H^2_M are measures of the relative magnitude of the between-study variance τ^2 and within-study variance s^2. Alternatively, if $\tau^2 = \tau^2_{\mathrm{DL}}$, then I^2 and H^2_M are measures of how much the point estimate of Q exceeds (or otherwise) its d.f. I^2 is the more well known of the two and is readily interpretable on the percentage scale. However, it is a nonlinear function of τ^2, such that an increase from, say, 80% to 90% implies a far greater increase in τ^2 than an increase from 20% to 30%. H^2_M does not have this limitation and is also less affected by the value of k, but it is perhaps less readily interpretable.

2.2   Random-effects methods

In this section, I describe the random-effects models currently available within the `ipdmetan` command.

Noniterative τ^2 estimators

The noniterative formulas for τ^2 presented in table 1 are based on moment-based approximations to the expectation of Q (DerSimonian–Laird; Hedges), a reparameterization of the total variance of θ_i (Sidik–Jonkman), or Bayesian considerations (Rukhin). Estimates of τ^2 are typically truncated at zero, but the Sidik–Jonkman and Rukhin BP estimators are always positive.

Table 1. Noniterative τ^2 estimators

Name	Reference	Definition
DerSimonian–Laird	DerSimonian and Laird (1986)	$$\tau^2_{\text{DL}} = \frac{Q - (k-1)}{\sum_i w_i - \frac{\sum_i w_i^2}{\sum_i w_i}}$$
Hedges "variance component" Cochran	Sidik and Jonkman (2007) DerSimonian and Kacker (2007)	$$\tau^2_{\text{VC}} = \frac{\sum_i \left(\theta_i - \overline{\theta}\right)^2}{k-1} - \frac{\sum_i \sigma_i^2}{k}$$ where $\overline{\theta} = \sum_i \theta_i / k$
Sidik–Jonkman	Sidik and Jonkman (2007)	$$\tau^2_{\text{SJ}} = \frac{\sum_i \widetilde{w_i}\left(\theta_i - \widetilde{\theta}\right)}{k-1}$$ where $\widetilde{w_i} = \{(\sigma_i^2)/(\tau_0^2) + 1\}^{-1}$ are inverse-variance weights where the variance estimator is based on a parameterization involving ratios of the within-study variances σ_i^2 to a crude initial estimate $\tau_0^2 > 0$ of the true heterogeneity variance, and $\widetilde{\theta}$ is the pooled estimate based on weights $\widetilde{w_i}$
Rukhin Bayes estimators	Rukhin (2013)	$$\tau^2_{\text{B0}} = \frac{\sum_i \left(\theta_i - \overline{\theta}\right)^2}{k+1} - \frac{(N-k)(k-1)\sum_i \sigma_i^2}{k(k+1)(N-k+2)}$$ where N is the total number of participants within the k included studies $$\tau^2_{\text{BP}} = \frac{\sum_i \left(\theta_i - \overline{\theta}\right)^2}{k+1}$$

The Sidik–Jonkman estimator requires a nonzero initial estimate τ_0^2, for which τ_{VC}^2 was suggested by Sidik and Jonkman (2007), with an arbitrary small value of 0.01 being used in case $\tau_{\mathrm{VC}}^2 = 0$. However, τ^2 is not scale invariant, so it is arguably more appropriate to assume an arbitrarily small I^2 instead, such as 1% (see the `isq()` suboption to `re()` in section 3.2).

Iterative τ^2 estimators

Iterative methods can be used to estimate τ^2 with confidence limits and, in the case of profile likelihood (PL), iterative confidence limits for θ. In table 2, w_i^* and $\widehat{\theta}^*$ are defined in terms of the current iteration of τ^2.

Table 2. Iterative τ^2 estimators

Name	Reference	Definition
Approximate gamma	Biggerstaff and Tweedie (1997)	Point estimate $\tau^2 = \tau^2_{\mathrm{DL}}$ Confidence limits for τ^2 obtained by profiling an approximate gamma distribution for Q Point estimate and standard error for θ obtained from modified weights generated by numerical integration
Mandel–Paule Generalized Q Empirical Bayes	DerSimonian and Kacker (2007) Viechtbauer (2007) Sidik and Jonkman (2007)	$$\tau^2_{\mathrm{GQ}} = \frac{k}{C}\frac{\sum_i w_i^* \left(\theta_i - \widehat{\theta}^*\right)^2}{\sum_i w_i^*} - \frac{\sum_i w_i^* \sigma_i^2}{\sum_i w_i^*}$$ where C is a critical value, as follows: $\quad$Point estimate $\qquad\qquad C = k - 1$ $\quad$Lower confidence limit $\quad C = \chi^2_{k-1;1-\alpha/2}$ $\quad$Upper confidence limit $\quad C = \chi^2_{k-1;\alpha/2}$
Maximum likelihood (ML)	Viechtbauer (2007)	$$\tau^2_{\mathrm{ML}} = \frac{\sum_i w_i^{*2}\left(\theta_i - \widehat{\theta}^*\right)^2}{\sum_i w_i^{*2}} - \frac{\sum_i w_i^{*2}\sigma_i^2}{\sum_i w_i^{*2}}$$ Confidence limits for τ^2 obtained by likelihood profiling
PL	Hardy and Thompson (1996)	τ^2 same as for ML. Confidence limits for θ obtained by nested likelihood profiling
Restricted maximum-likelihood (REML)	Viechtbauer (2007)	$$\tau^2_{\mathrm{REML}} = \frac{\sum_i w_i^{*2}\left(\theta_i - \widehat{\theta}^*\right)^2}{\sum_i w_i^{*2}}$$ $$-\frac{\sum_i w_i^{*2}\sigma_i^2}{\sum_i w_i^{*2}} + \frac{1}{\sum_i w_i^*}$$ Confidence limits for τ^2 obtained by likelihood profiling

Extensions to the DerSimonian–Laird model

    a. Hartung–Knapp variance estimator

      This approach, proposed by Hartung and Knapp (2001) and independently by Sidik and Jonkman (2002), uses an alternative estimator of $\text{Var}(\widehat{\theta})$ to give improved coverage probability in the presence of between-study heterogeneity rather than relying on accurate estimation of τ^2. Given standard DerSimonian–Laird random-effects weights $w_i^* = 1/(\sigma_i^2 + \tau_{\text{DL}}^2)$ and corresponding pooled effect $\widehat{\theta^*}$, the Hartung–Knapp variance estimator is defined as

$$\text{Var}\left(\widehat{\theta^*}\right) = \frac{\sum\limits_i w_i^* \left(\theta_i - \widehat{\theta^*}\right)^2}{(k-1)\sum\limits_i w_i^*}$$

      This is used to construct the following t-based $(1 - \alpha)$ confidence interval for $\widehat{\theta^*}$:

$$\widehat{\theta^*} \pm t_{k-1;1-\alpha/2}\sqrt{\text{Var}\left(\widehat{\theta^*}\right)}$$

    b. Nonparametric bootstrap

      Kontopantelis, Springate, and Reeves (2013) more recently suggested a simple nonparametric bootstrap version of DerSimonian–Laird by sampling studies with replacement and taking the mean of the truncated τ_{DL}^2 estimates. This is intended to increase the chance of detecting heterogeneity when the true value of τ^2 is small.

Various simulation studies have been performed to assess the properties of the estimators of θ and τ^2 described above. Kontopantelis and Reeves (2012) concluded that the PL and Hartung–Knapp methods gave the most accurate coverage probabilities for θ, even when the true study effects were extremely nonnormally distributed. However, confidence intervals may be too wide for small numbers of studies k; in this situation, they recommend the standard DerSimonian–Laird method. Rukhin (2013) and Kontopantelis, Springate, and Reeves (2013) found that the Bayes estimator τ_{BP}^2 gave reliable coverage and accurate confidence intervals for small k, except when the true heterogeneity was also small, in which case the bootstrapped DerSimonian–Laird estimator was preferable.

    In terms of τ^2, Sidik and Jonkman (2007) concluded that τ_{GQ}^2 and τ_{SJ}^2 were the least biased point estimates of τ^2. Viechtbauer (2007) found that the generalized Q and approximate gamma models gave the most accurate confidence intervals for τ^2.

    Kontopantelis, Springate, and Reeves (2013) examined real-life meta-analysis data from the Cochrane Collaboration and concluded that heterogeneity is often underestimated, especially for small k. They therefore propose that sensitivity analyses be performed at specified levels of heterogeneity. Such a level might be moderate to high to test the sensitivity of conclusions regarding θ, or it may be minimal to simply force nonzero heterogeneity (see the `isq()` suboption to `re()` in section 3.2).

2.3 Interactions

What initially motivated the development of `ipdmetan` was unbiased estimation of patient-level treatment-covariate interactions in IPD meta-analysis. As discussed in Fisher et al. (2011), additional complexities exist for IPD that do not occur with AD. With AD, there is just one observation per study; hence, only "study-level" characteristics can be recorded—data such as year of publication, specifics of treatment, or study-average patient characteristics. Such data can be used only to investigate how characteristics of studies may affect study-aggregate treatment effects via techniques such as meta-regression and study-subgroup analysis. However, to investigate how treatment effects differ at the patient level requires IPD. An AD analysis of study-average patient characteristics (for example, mean age or proportion male) is at risk of bias as an estimate of the true patient-level interaction effect.

Patient-level characteristics are often categorical—for example, they may be categorized by sex, age subgroups, or disease stage. It may be tempting to compare by-subgroup treatment effects to estimate an interaction. However, this is again biased if interpreted as a measure of treatment-effect difference between patients (Fisher et al. 2011). Instead, treatment-covariate interactions must be estimated within each study separately and pooled thereafter in the same way as overall treatment effects. (Of course, continuous patient-level characteristics do not present such issues.)

Assume a series of linear regressions (extension to generalized linear models or time-to-event regressions is straightforward) on patients j within studies i as follows (the parentheses around i highlight the two-stage nature of the analysis):

$$y_{(i)j} = \alpha_{(i)} + \beta_{(i)}x_{(i)j} + \gamma_{(i)}z_{(i)j} + \delta_{(i)}x_{(i)j}z_{(i)j}$$

y is the outcome; x the treatment; z the covariate; and α, β, γ, and δ represent study, treatment, covariate, and interaction effect coefficients, respectively. We equate the interaction-effect estimates $\delta_{(i)}$ with the study estimates θ_i in section 2.1 and proceed accordingly (Simmonds and Higgins 2007; Fisher et al. 2011, *Appendix B*). Such models may be extended to include adjustment for (within-study) covariates.

Interaction models are often complex, and the two-stage approach is suitable for only simple cases—principally meta-analyses of RCTs where other factors may be assumed to be balanced. Furthermore, only one coefficient may be pooled, placing limitations on variable transformation (for example, use of fractional polynomials [Royston and Altman 1994]) and precluding the recovery of absolute effects. However, they are certain to be free of bias from between-study effects (Simmonds et al. 2005; Fisher et al. 2011, *Appendix B*) and to provide a quick and accurate estimate of the within-study interaction effect size. Although one-stage models permit a far greater variety of random effects, which may increase efficiency, the resulting coefficients must be interpreted carefully to ensure the intended question is answered.

3   The ipdmetan command

3.1   Syntax

ipdmetan is a prefix command; see [U] **11.1.10 Prefix commands**. Its basic syntax is

ipdmetan [*exp_list*], <u>stud</u>y(*study_id*[, <u>m</u>issing]) [by(*subgroup_id*[, <u>m</u>issing])
   eform_option <u>eff</u>ect(*string*) <u>inter</u>action keepall <u>mes</u>sages <u>nogra</u>ph
   <u>notab</u>le nohet nooverall <u>nosub</u>group <u>notot</u>al ovwt sgwt
   <u>poolv</u>ar(*model_coefficient*) random re[(*re_model*)] sortby(*varname* | _n)
   <u>lcols</u>(*cols_info*) <u>rcols</u>(*cols_info*) plotid(*varname* | _BYAD [, list <u>nogra</u>ph])
   ovstat(q) <u>saving</u>(*filename*[, replace <u>stack</u>label])
   <u>forest</u>plot(*forestplot_options*) ad(...)]:   *estimation_command* ...

where *estimation_command* is the model to be fit within each study. Typically, this will
be a built-in command that changes the contents of e(b), but it could be any command
that returns an effect size and standard error (or statistics that may be combined to
form such). if, in, and weights should be supplied to *estimation_command* rather than
to ipdmetan (that is, after the colon rather than before it). Relevant declarations of
data structure (for example, stset, tsset, and xtset) must be done before running
ipdmetan.

   When *estimation_command* does not change the contents of e(b), expressions that
evaluate the effect size and standard error to be collected and pooled must be specified
manually by using *exp_list*. (An optional third statistic is the number of observations.)
Otherwise, *exp_list* defaults to _b[*varname*] _se[*varname*], where *varname* is the first
independent variable within *estimation_command*.

3.2   Options

The options in this section (supplied directly to ipdmetan) control aspects of model
fitting, computation, creation of the results set, on-screen tabulation of results, and
saved statistics. Some options included are relevant to the forest plot, but they are used
within the main routine of ipdmetan. The only suboptions that should be supplied to
the forestplot() option are those controlling graphical presentation of the results set
outputted by ipdmetan.

study(*study_id*[, missing]) specifies the variable that contains the study identifier,
   which must be either integer valued or string. study() is required.

   missing requests that missing values be treated as potential study identifiers (the
     default is to exclude them).

by(*subgroup_id*[, `missing`]) specifies a variable identifying subgroups of studies (and must therefore be constant within studies), which must be integer valued or string. If an AD dataset is specified and contains a variable named *subgroup_id*, subgrouping will be extended to the AD observations.

> `missing` requests that missing values be treated as potential subgroup identifiers (the default is to exclude them).

eform_option (see [R] ***eform_option***) specifies that effect sizes and confidence limits be exponentiated in the output (table and forest plot) and generates a heading for the effect-size column.

> Note that `ipdmetan` does not check the validity of the particular *eform_option*; for example, it does not check whether *estimation_command* is a survival model if `hr` is supplied.

`effect(`*string*`)` specifies a heading for the effect-size column, overriding any heading generated by *eform_option*.

`interaction` specifies that *estimation_command* contain one or more interaction effects expressed using factor-variable syntax (see [U] **11.4.3 Factor variables**) and that the first valid interaction effect be pooled across studies. This is a helpful shortcut for simple interaction analyses, but it is not foolproof or comprehensive. The output should be checked and, if necessary, the analysis rerun with the `poolvar()` option.

`keepall` specifies that all values of *study_id* be visible in the output (table and forest plot), even if no effect could be estimated (for example, because of insufficient observations or missing data). For such studies, (`Insufficient data`) will appear in place of effect estimates and weights.

`messages` requests that information be printed to screen whether the effect size and standard-error statistics have been successfully obtained from each study and, if applicable, whether the iterative random-effects calculations converged successfully.

`nograph` and `notable` suppress construction of the forest plot and the table of effect sizes, respectively.

`nohet` suppresses all heterogeneity statistics.

`nooverall` suppresses the overall pooled effect so that, for instance, subgroups are considered independently. It also suppresses between-subgroup heterogeneity statistics (if applicable).

`nosubgroup` suppresses the within-subgroup pooled effects (if applicable) so that subgroups are displayed separately but with a single overall pooled effect with associated heterogeneity statistics.

`nototal` requests that *estimation_command* not be fit within the entire dataset, for example, for time-saving reasons. By default, such fitting is done to check for problems in convergence and in the validity of requested coefficients and returned expressions. If `nototal` is specified, either `poolvar()` or *exp_list* must be supplied, and a message

appears above the table of results warning that estimates should be double checked by the user.

ovwt and `sgwt` override the default choice of whether to display overall weights or within-subgroup weights in the screen output and forest plot. Note that because weights are normalized, these options do not affect estimation of pooled effects or heterogeneity statistics.

poolvar(*model_coefficient*) specifies the coefficient from `e(b)` that is to be pooled when the default behavior of `ipdmetan` fails or is incorrect. *model_coefficient* should be a variable name, a level indicator, an interaction indicator, or an interaction involving continuous variables (c.f. syntax of [R] **test**). Equation names can be specified using the format poolvar(*eqname*:*varname*). This option is appropriate only when *estimation_command* changes the contents of `e(b)`; otherwise, see *exp_list*.

random or `re` specifies DerSimonian and Laird random effects.

re(*re_model*) specifies other possible random-effects models. The default is the DerSimonian and Laird random-effects model. Other currently supported models (see section 2.2) are the following:

> dl specifies the DerSimonian–Laird estimator (equivalent to specifying `re` alone, with no suboption); this is the default.

> dlt or `hk` specifies the DerSimonian–Laird with Hartung–Knapp t-based variance estimator.

> bdl or `dlb` specifies the bootstrapped DerSimonian–Laird estimator.

> ca, `he`, or `vc` specifies the Hedges variance-component estimator, also known as the Cochran estimator.

> sj specifies the Sidik–Jonkman two-step estimator.

> b0 and `bp` specify Rukhin's $B0$ and BP estimators, respectively.

> bs, `bt`, or `gamma` specifies the Biggerstaff–Tweedie approximate gamma model.

> eb, `gq`, `genq`, `mp`, or `q` specifies the empirical Bayes estimator, also known as the generalized Q estimator and the Mandel–Paule estimator.

> ml specifies the ML estimator.

> pl specifies the PL model.

> reml specifies the REML estimator.

> sa$\big[$, isq(*real*)$\big]$ specifies the sensitivity analysis with a fixed user-specified value of I^2 (between 0 and 1, with default isq(0.8)). See Kontopantelis, Springate, and Reeves (2013).

Note that the approximate gamma, generalized Q, ML, PL, and REML models require the `mm_root()` function; and the bootstrapped DerSimonian–Laird model requires the `mm_bs()` and `mm_jk()` functions from the `moremata` package (`ssc install moremata`) (Jann 2005). The approximate gamma model also requires the `integrate` command (`ssc install integrate`) (Mander 2012).

`sortby(`*varname* | `_n)` allows user-specified ordering of studies in the table and forest plot. The default ordering is by *study_id*. Note that `sortby()` does not alter the data in memory.

Specify `sortby(_n)` to order the studies by their first appearance in the data by using the current sort order.

`lcols(`*cols_info*`)` and `rcols(`*cols_info*`)` define columns of additional data to be presented to the left or right of the forest plot. These options are carried over from `metan`; however, for IPD, they must first be generated from the existing dataset and thus require more complex syntax.

cols_info has the following syntax, which is based on that of `collapse` (see [D] **collapse**),

$\big[\,(stat)\,\big]\,\big[\,newname\texttt{=}\big]item\,\big[\,\%fmt\,\texttt{"}label\texttt{"}\,\big]\,\big[\,\big[\,newname\texttt{=}\big]item\,\big[\,\%fmt\,\texttt{"}label\texttt{"}\,\big]\,\big]$ $\ldots\,\big[\,\big[\,(stat)\,\big]\,\ldots\,\big]$

where *stat* is as defined for `collapse`; *newname* is an optional user-specified variable name; *item* is either a numeric expression (in parentheses) involving returned quantities from *estimation_command* or a variable currently in memory; *%fmt* is an optional format; and "*label*" is an optional variable label. An example using these options is in section 4.5.

`plotid(`*varname* | `_BYAD` $\big[$ `, list nograph` $\big]$ `)` specifies one or more categorical variables to form a series of groups of observations in which specific aspects of plot rendition may be affected using $plot\big[\,\#\,\big]$`opts` (see *Forest plot suboptions*). The groups of observations will automatically be assigned ordinal labels $(1, 2, \ldots)$ on the basis of the ordering of *varlist*. Note that `plotid()` does not alter the data in memory.

`ovstat(q)` displays Q statistics instead of I-squared statistics.

`saving(`*filename*$\big[$ `, replace stacklabel` $\big]$ `)` saves the forest plot "results set" (see section 4.6) created by `ipdmetan` in a dataset for further use or manipulation.

`replace` overwrites *filename*.

`stacklabel` takes the variable label from the leftmost column variable (usually *study_id*), which would usually appear outside the plot region as the column heading, and copies it into a new first row in *filename*. This allows multiple such datasets to be `append`ed without this information being overwritten.

`forestplot(`*forestplot_options*`)` specifies other options to pass to `forestplot`. See *Forest plot suboptions* below.

The AD option

`ipdmetan` is designed so that IPD and AD can be analyzed together with a single command. Because of their differing structures (IPD is one observation per patient, so there are multiple observations per study; AD is one observation per study), AD must be stored in a separate Stata dataset.

The syntax for the `ad()` option is as follows:

`ad(`*filename* $\big[\,$*if*$\,\big]$ $\big[\,$*in*$\,\big]$`, vars(`*namelist*`)` *ad_options*`)`

where *filename* is an existing Stata dataset containing the AD.

`vars(`*namelist*`)` contains the names of variables (within *filename*) containing the effect size and either a standard error or lower and upper 95% confidence limits on the linear scale. If confidence limits are supplied, they must be derived from a normal distribution, or the pooled result will be incorrect (see help `admetan`).

It is assumed that a study identifier variable exists with the same name as in the data currently in memory (that is, the IPD). If such a variable cannot be found, studies in the AD dataset are numbered sequentially, following the largest study identifier value used for the IPD. Subgroup variables may also be present in the AD dataset and may be referenced using the main `by()` option.

ad_options can be the following:

`byad` specifies that IPD and AD be treated as subgroups rather than as a single set of estimates.

`npts(`*varname*`)` allows participant numbers (stored in *varname* within *filename*) to be displayed in tables and forest plots.

Forest plot suboptions

Most of the forest plot options for `metan` can be supplied to `ipdmetan` via the option `forestplot()`, together with most appropriate [G-3] *twoway_options*. The most significant extension that `ipdmetan` brings is the ability to apply specific plot rendition options (color, line patterns, etc.) to specific groups of observations via the `plotid()` option (see above).

The syntax for these options is

plot$\big[\,\#\,\big]$`opts(`*plot_options*`)`

where *plot* refers to the plot feature, which can be `box`, `ci`, `diam`, `oline`, `point`, `pci`, or `ppoint`; # is the (optional) observation group identifier assigned by `plotid()`; and *plot_options* are [G-3] *twoway_options* appropriate to the plot type used to construct the feature, as described in Harris et al. (2008).

The extra plot features `pci` and `ppoint`, not included in `metan`, are used when pooled estimates must be represented by (differently rendered) point estimates and confidence intervals instead of by diamonds. See the `forestplot` documentation for more details.

3.3 Stored results

`ipdmetan` stores the following in `r()` (with some variation):

Scalars

`r(k)`	number of included studies k	`r(tausq)`	between-study variance tau-squared
`r(n)`	number of included participants	`r(sigmasq)`	average within-study variance
`r(eff)`	overall pooled effect size	`r(Isq)`	heterogeneity measure I^2
`r(se_eff)`	standard error of pooled effect size	`r(HsqM)`	heterogeneity measure H_M^2
`r(Q)`	Cochran Q heterogeneity statistic (on $k-1$ d.f.)		

Macros

`r(command)`	full estimation command line	`r(estvar)`	name of pooled coefficient
`r(cmdname)`	estimation command name	`r(re_model)`	random-effects model used

Matrices

`r(coeffs)`	matrix of study and subgroup identifiers, effect coefficients, numbers of participants, and weights

Certain iterative random-effects models can save the following additional results (see `help mf_mm_root` for interpretations of convergence success values):

Scalars

`r(tsq_var)`	estimated variance of tau-squared
`r(tsq_lci)`	lower confidence limit for tau-squared
`r(tsq_uci)`	upper confidence limit for tau-squared
`r(rc_tausq)`	convergence of tau-squared point estimate
`r(rc_tsq_lci)`	convergence of tau-squared lower confidence limit
`r(rc_tsq_uci)`	convergence of tau-squared upper confidence limit
`r(rc_eff_lci)`	convergence of effect-size lower confidence limit
`r(rc_eff_uci)`	convergence of effect-size upper confidence limit

4 Example

The motivating example is an IPD meta-analysis of RCTs evaluating the effectiveness of postoperative radiotherapy in patients with completely resected nonsmall cell lung cancer (Burdett, Stewart, and PORT Meta-analysis Group 2005; Burdett et al. 2013). The outcome measure in each trial is overall survival, that is, censored time to death from any cause. A number of covariate measurements were also available but not necessarily for all trials or patients.

For an example dataset to be made publicly available, a simulated dataset was generated with similar characteristics to the postoperative radiotherapy data. It is this simulated dataset that will be discussed henceforth. There are 10 trials of varying size

in 2 trial subgroups with arbitrary names. Sex, age, and disease-stage covariates have also been generated.

4.1   Basic use

Our first example aims to cover various aspects of inverse-variance meta-analysis simultaneously, because none of this functionality is particularly complex. We perform a meta-analysis of main treatment effects fit from Cox models stratified by sex, pooling the log hazard-ratios using the fixed-effects inverse-variance method.

```
. use ipdmetan_example

. quietly stset tcens, fail(fail)

. ipdmetan, study(trialid) hr by(region)
> forest(favours(Favours treatment # Favours control)): stcox trt, strata(sex)

Studies included: 10
Patients included: 1642

Meta-analysis pooling of main (treatment) effect estimate trt
using Fixed-effects
```

Region and Trial name	Haz. Ratio	[95% Conf. Interval]		% Weight
Europe				
London	1.475	1.015	2.143	10.93
Paris	1.197	0.789	1.816	8.78
Amsterdam	1.742	1.080	2.810	6.68
Stockholm	1.440	0.816	2.542	4.73
Madrid	2.242	1.501	3.350	9.47
Subgroup effect	1.593	1.312	1.934	40.59
North America				
New York	0.740	0.465	1.176	7.10
Chicago	0.774	0.558	1.074	14.29
Los Angeles	1.058	0.707	1.584	9.38
Toronto	1.082	0.816	1.434	19.21
College Station, TX	0.689	0.461	1.031	9.42
Subgroup effect	0.885	0.754	1.039	59.41
Overall effect	1.123	0.993	1.271	100.00

```
Tests of effect size = 1:
Europe                z =   4.707  p =   0.000
North America         z =  -1.495  p =   0.135
Overall               z =   1.846  p =   0.065

Q statistics for heterogeneity (calculated using Inverse Variance weights)
```

	Value	df	p-value
Europe	5.01	4	0.287
North America	5.40	4	0.249
Overall	31.39	9	0.000
Between	20.98	1	0.000
Between:Within (F)	6.02	1, 9	0.037

The command `stcox trt, strata(sex)` was first fit consecutively to each level of `trialid`. The resulting log hazard-ratios and standard errors `_b[trt]` and `_se[trt]` are displayed in the table and forest plot (figure 1). Specifying the *eform_option* `hr` exponentiated the estimates into hazard ratios, while the `by()` option gave us within-subgroup pooled estimates and tests of heterogeneity within and between subgroups and overall. The form of the rest of the output and the forest plot should be familiar to users of `metan` (Harris et al. 2008).
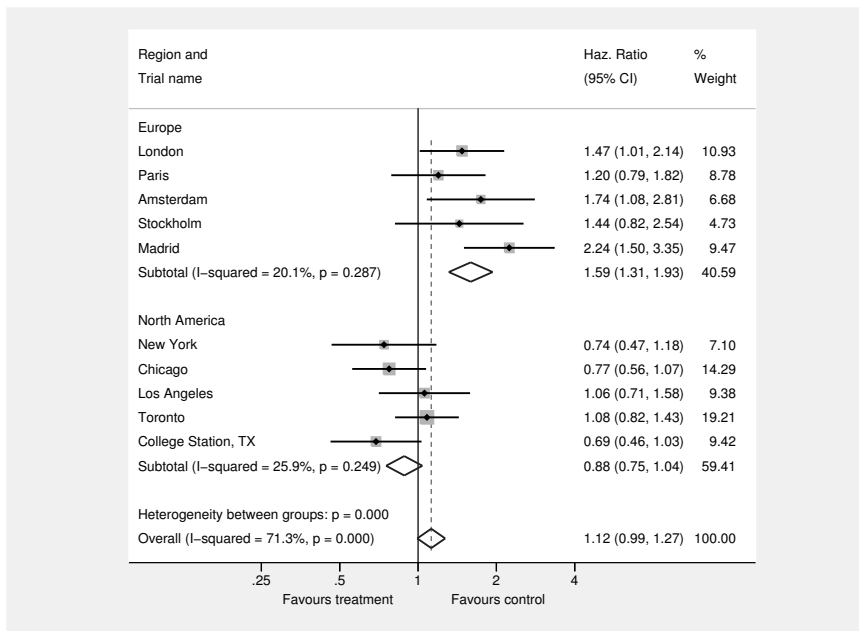

Figure 1. Basic forest plot

4.2   Treatment-covariate interactions

The syntax of `ipdmetan` allows any (single) estimated quantity and its standard error to be pooled across studies. This includes simple interaction models of the form given in section 2.3. Specifying the `interaction` option tells `ipdmetan` to identify and pool the treatment-covariate interaction coefficient and to use different symbols and effect-size column headings on the forest plot to differentiate such analyses from those of main effects. (Note that all of this behavior can be reproduced manually using other options.)

Studies often have insufficient covariate data, and such studies will be excluded from the corresponding interaction analysis. The default is to remove excluded studies from the output. However, this can be overridden with the `keepall` option, which will instead output "`(Insufficient data)`" for such studies in place of the effect size, the standard error, and the weighting in both table and forest plot. (This option applies generally. I mention it here because it is likely to have greater relevance.)

The screen output has only minor differences from that of a main-effects analysis, so we show only the forest plot here (figure 2). The default for interaction analyses is to use solid circles for the study effects and a clear circle for the pooled effect.

```
. ipdmetan, study(trialid) interaction hr keepall forest(favours("Favours
> greater treatment effect" "with higher disease stage"
> # "Favours greater treatment effect" "with lower disease stage")
> boxsca(200) fp(1)): stcox trt##c.stage
```
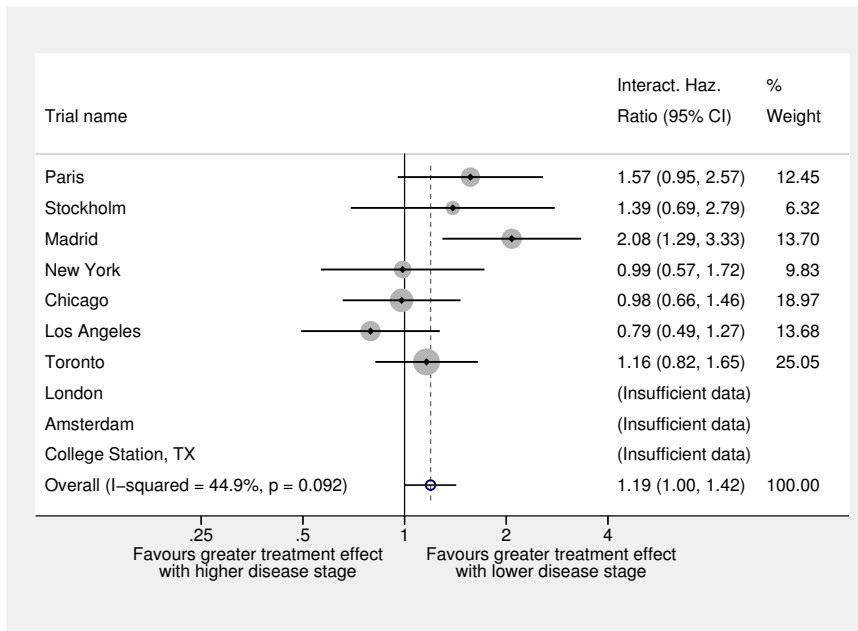

Figure 2. Forest plot of interactions

4.3   Random effects

We now show the syntax and output of the random-effects options. DerSimonian–Laird, the most common random-effects model, is specified with "`re`".

▷ **Example 1: DerSimonian–Laird random effects**

```
. ipdmetan, study(trialid) nograph re: stcox trt
    (output omitted)
Heterogeneity Measures
```

	Value	df	p-value
Cochran Q	31.73	9	0.000
I^2 (%)	71.6%		
Modified H^2	2.525		
tau^2	0.0985		

```
I² = between-study variance (tau²) as a percentage of total variance
Modified H² = ratio of tau² to typical within-study variance
```

◁

Other random-effects models are specified as options to "`re`", as shown in the following example (see section 2.2 for details of the Generalized Q model). Note that iterative models output confidence limits for heterogeneity in addition to point estimates.

▷ **Example 2: Generalized Q random effects**

```
. ipdmetan, study(trialid) nograph re(q): stcox trt
    (output omitted)
Heterogeneity Measures
```

	Value	df	p-value
Cochran Q	31.73	9	0.000

	Value	[% Conf. Interval]	
I^2 (%)	82.7%	73.8%	93.1%
Modified H^2	4.779	2.823	13.390
tau^2	0.1864	0.1101	0.5223

```
I² = between-study variance (tau²) as a percentage of total variance
Modified H² = ratio of tau² to typical within-study variance
```

◁

4.4   Aggregate data

Often IPD cannot be obtained for all eligible studies. In such cases, it may help to do a sensitivity analysis comparing "extra" AD estimates with the main body of IPD. Some authors (Riley et al. 2008) have even suggested pooling both IPD and AD together. Either of these approaches is straightforward with `ipdmetan`. Here we demonstrate the former by first artificially constructing an aggregate dataset from one of the trial subgroups we defined earlier by using the `saving()` option. Note the following (see also section 4.6):

- Saved datasets have a standard format where the study variable is named _STUDY. Therefore, we change our `trialid` variable to match (using `clonevar`) and so obtain correct labels throughout.

- We restrict the `ad()` option to _USE==1 so that only individual study estimates are included and not the pooled estimate.

```
. quietly ipdmetan, study(trialid) hr nograph saving(region2.dta): stcox
> trt if region==2, strata(sex)
. clonevar _STUDY = trialid
. ipdmetan, study(_STUDY) hr ad(region2.dta if _USE==1, vars(_ES _seES)
> npts(_NN) byad) nooverall: stcox trt if region==1, strata(sex)
Studies included from IPD: 5
Patients included: 656

Studies included from aggregate data: 5
Patients included: 986

Meta-analysis pooling of main (treatment) effect estimate trt
using Fixed-effects
```

Data source and Trial name	Haz. Ratio	[95% Conf. Interval]		% Weight
IPD				
London	1.475	1.015	2.143	26.93
Paris	1.197	0.789	1.816	21.63
Amsterdam	1.742	1.080	2.810	16.46
Stockholm	1.440	0.816	2.542	11.64
Madrid	2.242	1.501	3.350	23.34
Subgroup effect	1.593	1.312	1.934	100.00
Aggregate				
New York	0.740	0.465	1.176	11.96
Chicago	0.774	0.558	1.074	24.06
Los Angeles	1.058	0.707	1.584	15.79
Toronto	1.082	0.816	1.434	32.33
College Station, TX	0.689	0.461	1.031	15.86
Subgroup effect	0.885	0.754	1.039	100.00

```
Tests of effect size = 1:
IPD                 z =   4.707  p =   0.000
Aggregate           z =  -1.495  p =   0.135

Q statistics for heterogeneity (calculated using Inverse Variance weights)
```

	Value	df	p-value
IPD	5.01	4	0.287
Aggregate	5.40	4	0.249

The numerical output is the same as in our first example, as expected. Note that to treat the IPD and AD as two subgroups, we specified the `byad` option; this also labels the subgroups appropriately in both the table and forest plot (not shown). The `nooverall`

option requests that the table and forest plot present IPD and AD separately and not present an "overall" pooled result. This will also present weights and effect lines by subgroup rather than overall. Above the table, the number of trials and patients are also given separately for IPD and for AD. If the number of patients is missing for one or more observations in the AD dataset, `ipdmetan` will output a missing value here.

4.5   Further functionality

In this section, I discuss a few more useful features of `ipdmetan` and `forestplot`. For this next example, we return to the original analysis in section 4.1, but instead of Cox regression, we will use the noniterative Peto log-rank method. This is based on the log-rank test done by `sts test`, but it requires values stored in the vector **u** and matrix **V** (see documentation of the `mat()` option in [ST] **sts test**) to estimate the effect size (log hazard-ratio) and standard error for each trial. Therefore, we must supply `ipdmetan` with appropriate expressions for these quantities as an *exp_list*. Using our current example, we also demonstrate the use of `lcols()` and `rcols()` by presenting the $O - E$ and V statistics (Yusuf et al. 1985), and we demonstrate the use of $plot[\,\#\,]$`opts` by rendering the weighted boxes and confidence interval lines in a light shade for the European trials (with capped confidence intervals) and in a darker shade for the North American trials. We proceed as follows (see figure 3):

```
. ipdmetan (u[1,1]/V[1,1]) (1/sqrt(V[1,1])), study(trialid) hr
> rcols((u[1,1]) %5.2f "o-E(o)" (V[1,1]) %5.1f "V(o)")
> by(region) plotid(region)
> forest(nooverall nostats nowt
> favours(Favours treatment # Favours control)
> box1opts(mcolor(gs10)) ci1opts(lcolor(gs10) rcap)
> box2opts(mcolor(gs6)) ci2opts(lcolor(gs6))): sts test trt, mat(u V)
```
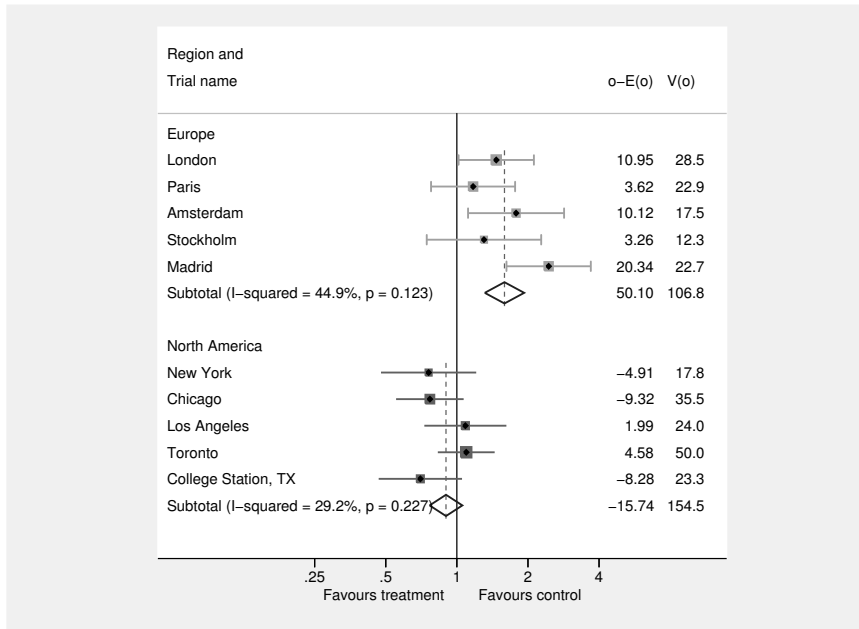
Figure 3. Use of advanced forest plot options

For the right-hand columns, we have requested that the usual statistics and weights be suppressed (options `nostats` and `nowt`) and replaced with the observed minus-expected event count and variance of observed event count (stored by `sts test` in `u[1,1]` and `V[1,1]`, respectively). Following each statistic is an (optional) format and column heading. (Note that this plot may be re-created with a Peto log-rank meta-analysis by using `petometan`, which is part of the `ipdmetan` package.)

Forest plots, with their columns of text and numbers, are unnatural to Stata, and sometimes, plots produced by `ipdmetan` or `forestplot` may be deficient in appearance (text too small or overlapping, too much white space, etc.). There are many options available that can correct such deficiencies (see the help file).

4.6   Use of results sets

Within a two-stage framework, IPD and AD may be considered a "long form" and "short form" of the same data. It may be useful to convert IPD to AD to perform certain tasks. This can be accomplished by using the `saving()` option with `ipdmetan` (see section 3.2), which saves an AD results set in a format suitable for `forestplot`. Hence, running `forestplot` on the results set will give the same plot as running `ipdmetan` on the original data.

As an example, the results set outputted by the analysis in section 4.1 is shown in the table below (value labels not displayed). This format includes not only the

study estimates but also the pooled estimates with extra lines for subgroup headings and heterogeneity descriptions and blank lines for spacing. The values of the variable _USE specify the function of each observation, while text (study names, headings, etc., excluding extra data specified in `lcols()` and `rcols()`) is stored in the variable _LABELS.

Possible uses of such results sets include the following:

- Restricting the results set to observations for which _USE==1 (that is, the individual study estimates) creates a standard AD dataset that can be analyzed with any AD meta-analysis command. Hence, features not yet supported by `ipdmetan` can be used, such as prediction intervals with `metan` (Harris et al. 2008) or the "permutations" model of `metaan` (Kontopantelis and Reeves 2010), as well as related analyses such as meta-regression on study-level covariates (`metareg`, Harbord and Higgins [2008]) or network meta-analysis (`mvmeta`, White [2009]).

- If the initial stage in `ipdmetan` of fitting the model to each study proves time consuming, it need be performed only once. Subsequent tasks, such as investigating random-effects models (see previous bullet) or tweaking the forest plot (see following bullet), can be done with `admetan` or `forestplot` using the results set.

- The results set contains all the information that `forestplot` needs to construct the plot. Therefore, the layout of the forest plot can be customized by altering the results set; for instance, it can be altered to amend labels or titles, add or remove lines, or add or format extra columns of data (to be specified with `lcols()` and `rcols()`). For example, strings of the form "n/N" can be presented, where n is the number of events and N the total number of patients.

- More complex forest plots may not be possible to create with a single call to `ipdmetan`. Instead, `ipdmetan` might be run on a series of different datasets and results saved. The standard formatting of results sets allows straightforward appending to a single dataset using `append` (see [D] **append**), potentially with the `generate` suboption to create a variable to pass to the `plotid()` suboption of `forestplot()` (see section 4.5). It may then be necessary to alter labels or titles (see previous bullet).

```
. list, clean noobs nolabel
```

_USE	_BY	_STUDY	_LABELS	_ES	_seES	_LCI	_UCI	_WT	_NN
0	1	.	Europe	.	.	.	.	.	.
1	1	1	London	0.389	0.191	0.015	0.762	0.109	176
1	1	2	Paris	0.180	0.213	-0.237	0.597	0.088	141
1	1	3	Amsterdam	0.555	0.244	0.077	1.033	0.067	110
1	1	4	Stockholm	0.365	0.290	-0.204	0.933	0.047	69
1	1	5	Madrid	0.807	0.205	0.406	1.209	0.095	160
3	1	.	Subtotal (I-squared = 20.1%, p = 0.287)	0.466	0.099	0.272	0.660	0.406	656
6	1	.	.	.	.	.	.	.	.
0	2	.	North America	.	.	.	.	.	.
1	2	6	New York	-0.301	0.237	-0.765	0.162	0.071	115
1	2	7	Chicago	-0.256	0.167	-0.583	0.071	0.143	238
1	2	8	Los Angeles	0.056	0.206	-0.347	0.460	0.094	148
1	2	9	Toronto	0.079	0.144	-0.203	0.361	0.192	316
1	2	10	College Station, TX	-0.372	0.205	-0.775	0.030	0.094	169
3	2	.	Subtotal (I-squared = 25.9%, p = 0.249)	-0.122	0.082	-0.283	0.038	0.594	986
6	2	.	.	.	.	.	.	.	.
4	.	.	Heterogeneity between groups: p = 0.000	.	.	.	.	.	.
5	.	.	Overall (I-squared = 71.3%, p = 0.000)	0.116	0.063	-0.007	0.240	1.000	1642

5   Discussion

Meta-analysis is a widely used and constantly evolving technique that has made an enormous contribution to evidence-based policy and practice, particularly in health care. Although one-stage hierarchical models are increasingly seen and may have statistical advantages in some situations, the two-stage approach remains commonplace when IPD are available.

The `ipdmetan` command is a tool for performing two-stage IPD analysis simply and straightforwardly; it gives a standardized output and publication-standard forest plots (based on the original forest plot subroutine within `metan`). The syntax is designed to encompass varying levels of complexity, and the dedicated `forestplot` command allows great flexibility of presentation. The package also provides useful capabilities such as trial subgroup plots (via `ipdover`) and inclusion of AD.

Future work might include the ability to present prognostic effects in a forest plot with respect to a reference, as well as the ability to use multiple imputation for missing covariates using `mi`.

6   Acknowledgments

7   References

Biggerstaff, B. J., and R. L. Tweedie. 1997. Incorporating variability in estimates of heterogeneity in the random effects model in meta-analysis. *Statistics in Medicine* 16: 753–768.

Bowden, J., J. F. Tierney, M. Simmonds, A. J. Copas, and J. P. T. Higgins. 2011. Individual patient data meta-analysis of time-to-event outcomes: One-stage versus two-stage approaches for estimating the hazard ratio under a random effects model. *Research Synthesis Methods* 2: 150–162.

Burdett, S., L. Rydzewska, J. F. Tierney, D. J. Fisher, and PORT Meta-analysis Trialist Group. 2013. A closer look at the effects of postoperative radiotherapy by stage and nodal status: Updated results of an individual participant data meta-analysis in non-small-cell lung cancer. *Lung Cancer* 80: 350–352.

Burdett, S., L. Stewart, and PORT Meta-analysis Group. 2005. Postoperative radiotherapy in non-small-cell lung cancer: Update of an individual patient data meta-analysis. *Lung Cancer* 47: 81–83.

Crowther, M. J., M. P. Look, and R. D. Riley. 2014. Multilevel mixed effects parametric survival models using adaptive Gauss–Hermite quadrature with application to recurrent events and individual participant data meta-analysis. *Statistics in Medicine* 33: 3844–3858.

Deeks, J. J., D. G. Altman, and M. J. Bradburn. 2001. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In *Systematic Reviews in Health Care: Meta-analysis in Context*, ed. M. Egger, G. Davey Smith, and D. G. Altman, 2nd ed., 285–312. London: BMJ Books.

DerSimonian, R., and R. Kacker. 2007. Random-effects model for meta-analysis of clinical trials: An update. *Contemporary Clinical Trials* 28: 105–114.

DerSimonian, R., and N. Laird. 1986. Meta-analysis in clinical trials. *Controlled Clinical Trials* 7: 177–188.

Fisher, D. J., A. J. Copas, J. F. Tierney, and M. K. B. Parmar. 2011. A critical review of methods for the assessment of patient-level interactions in individual participant data meta-analysis of randomized trials, and guidance for practitioners. *Journal of Clinical Epidemiology* 64: 949–967.

Harbord, R. M., and J. P. T. Higgins. 2008. Meta-regression in Stata. *Stata Journal* 8: 493–519.

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

Harris, R. J., M. J. Bradburn, J. J. Deeks, R. M. Harbord, D. G. Altman, and J. A. C. Sterne. 2008. metan: Fixed- and random-effects meta-analysis. *Stata Journal* 8: 3–28.

Hartung, J., and G. Knapp. 2001. A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Statistics in Medicine* 20: 3875–3889.

Higgins, J. P. T., and S. Green. 2011. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. http://www.cochrane-handbook.org.

Higgins, J. P. T., and S. G. Thompson. 2002. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 21: 1539–1558.

Jann, B. 2005. moremata: Stata module (Mata) to provide various functions. Statistical Software Components S455001, Department of Economics, Boston College. https://ideas.repec.org/c/boc/bocode/s455001.html.

Kontopantelis, E., and D. Reeves. 2010. metaan: Random-effects meta-analysis. *Stata Journal* 10: 395–407.

———. 2012. Performance of statistical methods for meta-analysis when true study effects are non-normally distributed: A simulation study. *Statistical Methods in Medical Research* 21: 409–426.

———. 2013. A short guide and a forest plot command (ipdforest) for one-stage meta-analysis. *Stata Journal* 13: 574–587.

Kontopantelis, E., D. A. Springate, and D. Reeves. 2013. A re-analysis of the Cochrane Library data: The dangers of unobserved heterogeneity in meta-analyses. *PLOS ONE* 8: e69930.

Mander, A. 2012. integrate: Stata module to perform one-dimensional integration. Statistical Software Components S457429, Department of Economics, Boston College. https://ideas.repec.org/c/boc/bocode/s457429.html.

Mathew, T., and K. Nordström. 2010. Comparison of one-step and two-step meta-analysis models using individual patient data. *Biometrical Journal* 52: 271–287.

Mittlböck, M., and H. Heinzl. 2006. A simulation study comparing properties of heterogeneity measures in meta-analyses. *Statistics in Medicine* 25: 4321–4333.

Riley, R. D., P. C. Lambert, J. A. Staessen, J. Wang, F. Gueyffier, L. Thijs, and F. Boutitie. 2008. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Statistics in Medicine* 27: 1870–1893.

Royston, P., and D. G. Altman. 1994. Regression using fractional polynomials of continuous covariates: Parsimonious parametric modelling (with discussion). *Journal of the Royal Statistical Society, Series C* 43: 429–467.

Rukhin, A. L. 2013. Estimating heterogeneity variance in meta-analysis. *Journal of the Royal Statistical Society, Series B* 75: 451–469.

Sidik, K., and J. N. Jonkman. 2002. A simple confidence interval for meta-analysis. *Statistics in Medicine* 21: 3153–3159.

———. 2007. A comparison of heterogeneity variance estimators in combining results of studies. *Statistics in Medicine* 26: 1964–1981.

Simmonds, M. C., and J. P. T. Higgins. 2007. Covariate heterogeneity in meta-analysis: Criteria for deciding between meta-regression and individual patient data. *Statistics in Medicine* 26: 2982–2999.

Simmonds, M. C., J. P. T. Higgins, L. A. Stewart, J. F. Tierney, M. J. Clarke, and S. G. Thompson. 2005. Meta-analysis of individual patient data from randomized trials: A review of methods used in practice. *Clinical Trials* 2: 209–217.

Sterne, J. A. C., ed. 2009. *Meta-Analysis in Stata: An Updated Collection from the Stata Journal.* College Station, TX: Stata Press.

Stewart, G. B., D. G. Altman, L. M. Askie, L. Duley, M. C. Simmonds, and L. A. Stewart. 2012. Statistical analysis of individual participant data meta-analyses: A comparison of methods and recommendations for practice. *PLOS ONE* 7: e46042.

Stewart, L. A., and J. F. Tierney. 2002. To IPD or not to IPD? Advantages and dis-
    advantages of systematic reviews using individual patient data. *Evaluation and the
    Health Professions* 25: 76–97.

Tudur Smith, C., and P. R. Williamson. 2007. A comparison of methods for fixed effects
    meta-analysis of individual patient data with time to event outcomes. *Clinical Trials*
    4: 621–630.

Viechtbauer, W. 2007. Confidence intervals for the amount of heterogeneity in meta-
    analysis. *Statistics in Medicine* 26: 37–52.

White, I. R. 2009. Multivariate random-effects meta-analysis. *Stata Journal* 9: 40–56.

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.

About the author

David Fisher is a statistician at the MRC Clinical Trials Unit at University College London
in London. He currently divides his time between clinical trials in colorectal cancer and the
MRC Clinical Trials Unit Meta-analysis Group, which carries out systematic reviews and IPD
meta-analyses as well as practical methodology research.