The Stata Journal (2009)
9, Number 1, pp. 40–56

Multivariate random-effects meta-analysis

Ian R. White
MRC Biostatistics Unit
Cambridge, UK
ian.white@mrc-bsu.cam.ac.uk

Abstract. Multivariate meta-analysis combines estimates of several related parameters over several studies. These parameters can, for example, refer to multiple outcomes or comparisons between more than two groups. A new Stata command, `mvmeta`, performs maximum likelihood, restricted maximum likelihood, or method-of-moments estimation of random-effects multivariate meta-analysis models. A utility command, `mvmeta_make`, facilitates the preparation of summary datasets from more detailed data. The commands are illustrated with data from the Fibrinogen Studies Collaboration, a meta-analysis of observational studies; I estimate the shape of the association between a quantitative exposure and disease events by grouping the quantitative exposure into several categories.

Keywords: st0156, mvmeta, mvmeta_make, mvmeta_l, meta-analysis, multivariate meta-analysis, individual participant data, observational studies

1 Introduction

Standard meta-analysis combines estimates of one parameter over several studies (Normand 1999). Multivariate meta-analysis is an extension that can combine estimates of several related parameters (van Houwelingen, Arends, and Stijnen 2003). In such work, it is important to allow for heterogeneity between studies, usually by fitting a random-effects model (Thompson 1994).

Multivariate meta-analysis has a variety of applications in randomized controlled trials. The simplest is modeling the outcome separately in each arm of a clinical trial (van Houwelingen, Arends, and Stijnen 2003). Other published applications explore treatment effects simultaneously on two clinical outcomes (Berkey, Anderson, and Hoaglin 1996; Berkey et al. 1998; Riley et al. 2007a,b) or on cost and effectiveness (Pinto, Willan, and O'Brien 2005), and explore combining trials comparing more than one treatment (Hasselblad 1998; Lu and Ades 2004). Further applications have been reviewed by Riley et al. (2007b).

There are also possible applications of multivariate meta-analysis in observational studies. These applications include assessing the shape of the association between a quantitative exposure and a disease, which will be illustrated in this article.

One difficulty in random-effects meta-analysis is estimating the between-studies variance. In the univariate case, this is commonly performed by using the method of DerSimonian and Laird (1986). However, maximum likelihood (ML) and restricted maximum likelihood (REML) methods are alternatives (van Houwelingen, Arends, and

Stijnen 2003); in Stata, they are not available in `metan` but can be obtained from `metareg` (Sharp 1998). This article describes a new command, `mvmeta`, that performs REML and ML estimation in the multivariate case by using a Newton–Raphson procedure. `mvmeta` requires a dataset of study-specific point estimates and their variance–covariance matrix. I also describe a utility command, `mvmeta_make`, that facilitates forming this dataset.

2 Multivariate random-effects meta-analysis with mvmeta

2.1 Syntax

`mvmeta` *b V* [*if*] [*in*] [`,` `reml` `ml` `mm` `fixed` `vars(`*varlist*`)` `corr(`*expression*`)`
`start(`*matrix*|*matrix_expression*|`mm)` `showstart` `showchol`
`keepmat(`*bname Vname*`)` `nouncertainv` `eform(`*name*`)` `bscorr` `bscov`
`missest(#)` `missvar(#)` *maximize_options*]

where the data are arranged with one line per study, the point estimates are held in variables whose names start with *b* (excluding *b* itself), the variance of *bx* is held in variable *Vxx*, and the covariance of *bx* and *by* is held in variable *Vxy* or *Vyx* (or the `corr()` option is specified).

If the dataset includes variables whose names start with *b* that do not represent point estimates, then the `vars()` option must be used.

2.2 Options

`reml`, the default, specifies that REML be used for estimation. Specify only one of the `reml`, `ml`, `mm`, or `fixed` options.

`ml` specifies that ML be used for estimation. ML is likely to underestimate the variance, so REML is usually preferred. Specify only one of the `reml`, `ml`, `mm`, or `fixed` options.

`mm` specifies that the multivariate method-of-moments procedure (Jackson, White, and Thompson Forthcoming) be used for estimation. This procedure is a multivariate generalization of the procedure of DerSimonian and Laird (1986) and is faster than the likelihood-based methods. Specify only one of the `reml`, `ml`, `mm`, or `fixed` options.

`fixed` specifies that the fixed-effects model be used for estimation. Specify only one of the `reml`, `ml`, `mm`, or `fixed` options.

`vars(`*varlist*`)` specifies which variables are to be used. By default, all variables *b** are used (excluding *b* itself). The order of variables in *varlist* does not affect the model itself but does affect the parameterization.

`corr(`*expression*`)` specifies that all within-study correlations take the given value. This means that covariance variable *Vxy* need not exist. (If it does exist, `corr()` is ignored.)

`start(`*matrix*`|`*matrix_expression*`|mm)` specifies a starting value for the between-studies variance, except `start(mm)` specifies that the starting value is computed by the `mm` method. If `start()` is not specified, the starting value is the weighted between-studies variance of the estimates, not allowing for the within-study variances; this ensures that the starting value is greater than zero (the iterative procedure never moves away from zero). `start(0)` uses a starting value of 0.001 times the default. The starting value for the between-studies mean is the fixed-effects estimate.

`showstart` reports the starting values used.

`showchol` reports the estimated values of the basic parameters underlying the between-studies variance matrix (the Cholesky decomposition).

`keepmat(`*bname Vname*`)` saves the vector of study-specific estimates and the vector of the variance–covariance matrix for study i as *bnamei* and *Vnamei*, respectively.

`nouncertainv` invokes alternative (smaller) standard errors that ignore the uncertainty in the estimated variance–covariance matrix and therefore agree with results produced by procedures such as SAS PROC MIXED (without the ddfm=kr option) and `metareg`. (Note, however, that the confidence intervals do not agree because `mvmeta` uses a normal approximation, whereas the other procedures approximate the degrees of freedom of a t distribution.)

`eform(`*name*`)` exponentiates the reported mean parameters, labeling them *name*.

`bscorr` reports the between-studies variance–covariance matrix as the standard deviations and reports the correlation matrix. This is the default if `bscov` is not specified.

`bscov` reports the between-studies variance–covariance matrix without transformation.

`missest(#)` specifies the value to be used for missing point estimates; the default is `missest(0)`. This is of minor importance because the variance of these missing estimates is specified to be very large.

`missvar(#)` is used in imputing the variance of missing point estimates. For a specific variable, the variance used is the largest observed variance multiplied by the specified value. The default is `missvar(1E4)`; this value is unlikely to need to be changed.

maximize_options are any options allowed by `ml maximize`.

3 Details of mvmeta

3.1 Notation

The data for `mvmeta` comprise the point estimate, y_i, and the within-study variance–covariance matrix, S_i, for each study $i = 1$ to n.

We assume the model

$$\begin{aligned} y_i &\sim N(\mu_i, S_i) \\ \mu_i &\sim N(\mu, \Sigma) \\ \Sigma &= \begin{pmatrix} \tau_1^2 & \kappa_{12}\tau_1\tau_2 & . \\ \kappa_{12}\tau_1\tau_2 & \tau_2^2 & . \\ . & . & . \end{pmatrix} \end{aligned}$$

where y_i, μ_i, and μ are $p \times 1$ vectors, and S_i and Σ are $p \times p$ matrices. The within-study variance, S_i, is assumed to be known. Our aim is to estimate μ and Σ.

We set $W_i = (\Sigma + S_i)^{-1}$, noting that this depends on the unknown Σ. If Σ were known (or assumed to be the zero matrix, as in fixed-effects meta-analysis), then we would have

$$\widehat{\mu} = \left(\sum_i W_i \right)^{-1} \left(\sum_i W_i y_i \right)$$

3.2 Estimating Σ

Methods proposed for estimating Σ in the multivariate setting include extensions of Cochran's method (Berkey et al. 1998), of the DerSimonian and Laird method (Pinto, Willan, and O'Brien 2005) for diagonal W_i, and of likelihood-based methods (van Houwelingen, Arends, and Stijnen 2003). We use the latter because of their generality and optimality properties. Respectively, the likelihood and restricted likelihood are

$$-2L = \sum_i \{\log|\Sigma + S_i| + (y_i - \mu)' W_i (y_i - \mu)\} + np\log 2\pi$$

$$-2RL = -2L + \log\left|\sum_i W_i\right| - p\log 2\pi \tag{1}$$

where W_i is a function of the unknown Σ, as noted above.

We maximize the (restricted) likelihood with a Newton–Raphson algorithm by using Stata's `ml` procedure. To ensure that Σ is nonnegative definite (for example, in the bivariate case, to ensure that the between-studies variances are nonnegative and that the between-studies correlation lies between -1 and 1), the basic model parameters are taken as the elements of a Cholesky decomposition of Σ (Riley et al. 2007b).

3.3 Saved results

As well as the usual `e()` information, `mvmeta` returns the estimated overall mean in `e(Mu)` and the between-studies variance–covariance matrix, the standard deviation vector, and the correlation matrix in `e(Sigma)`, `e(Sigma_SD)`, and `e(Sigma_corr)`, respectively.

3.4 Files required

`mvmeta` uses the likelihood program `mvmeta_l.ado`.

4 A utility command to produce data in the correct format: mvmeta_make

4.1 Syntax

`mvmeta_make` *regression_command* [*if*] [*in*] [*weight*], `by(`*by_variable*`)`
`saving(`*savefile*`)` [`replace append names(`*bname Vname*`) keepmat`
`usevars(`*varlist*`) useconstant esave(`*namelist*`) nodetails pause`
`ppfix(none|check|all) augwt(#) noauglist ppcmd(`*regcmd*[`,` *options*]`)`
`hard` *regression_options*]

`mvmeta_make` performs *regression_command* for each level of *by_variable* and stores the results in *savefile* in the format required by `mvmeta`. *weight* is any weight allowed by *regression_command*.

4.2 Options

`by(`*by_variable*`)` is required; it identifies the studies in which the regression command will be performed.

`saving(`*savefile*`)` is required; it specifies to save the regression results to *savefile*.

`replace` specifies to overwrite the existing file called *savefile*.

`append` specifies to append the current results to the existing file called *savefile*.

`names(`*bname Vname*`)` specifies that the estimated coefficients for variable x are to be stored in variable *bnamex* and that the estimated covariance between coefficients *bnamex* and *bnamey* is to be stored in variable *Vnamexy*. The default is `names(y S)`.

`keepmat` specifies that the results are also to be stored as matrices. The estimate vector and the covariance matrix for study i are stored as matrices *bnamei* and *Vnamei*, respectively, where *bname* and *Vname* are specified with `names()`.

`usevars(`*varlist*`)` identifies the variables whose regression coefficients are of interest. The default is all variables in the model, excluding the constant.

`useconstant` specifies that the constant is also of interest.

`esave(`*namelist*`)` adds the specified `e()` statistics to the saved data. For example, `esave(N ll)` saves `e(N)` and `e(ll)` as variables `_e_N` and `_e_ll`.

`nodetails` suppresses the results of running *regression_command* on each study.

`pause` pauses output after the analysis of each study, provided that `pause on` has been set.

`ppfix(none|check|all)` specifies whether perfect prediction should be fixed in no studies, only in studies where it is detected (the default), or in all studies.

`augwt(#)` specifies the total weight of augmented observations to be added in any study in which perfect prediction is detected (see section 7). `augwt(0)` turns off augmentation but is not recommended. The default is `augwt(0.01)`.

`noauglist` suppresses listing of the augmented observations.

`ppcmd(`*regcmd*`[ , `*options*`])` specifies that perfect prediction should be fixed by using regression command *regcmd* with options *options* instead of by using the default augmentation procedure.

`hard` is useful when convergence cannot be achieved in some studies. It captures the results of initial model fitting in each study and treats any nonzero return code as a symptom of perfect prediction.

regression_options are any options for *regression_command*.

5 Example 1: Telomerase data

Data from 10 studies of the value of telomerase measurements in the diagnosis of primary bladder cancer were reproduced by Riley et al. (2007b). In the table below, taken from that article, `y1` is logit sensitivity, `y2` is logit specificity, and `s1` and `s2` are their respective standard errors, all estimated from 2×2 tables of true status versus test status.

```
. use telomerase
(Riley´s telomerase data)
. format y1 s1 y2 s2 %6.3f
. list, noobs clean
    study      y1      s1       y2      s2
        1   1.139   0.406    3.219   1.020
        2   1.447   0.556    1.299   0.651
        3   1.705   0.272    0.661   0.308
        4   0.470   0.403    3.283   0.588
        5   0.856   0.290    4.920   1.004
        6   1.440   0.371    1.386   0.456
        7   0.187   0.306    3.219   1.442
        8   1.504   0.451    2.197   0.745
        9   1.540   0.636    2.269   0.606
       10   1.665   0.412   -1.145   0.434
. generate S11=s1^2
. generate S22=s2^2
```

5.1 Univariate meta-analysis

We first analyze the data by two univariate meta-analyses:

```
. mvmeta y S, vars(y1) bscov
Note: using method reml
Note: using variable y1
Note: 10 observations on 1 variables
  (output omitted)
                                                Number of obs   =         10
                                                Wald chi2(1)    =      38.52
Log likelihood = -8.7276382                     Prob > chi2     =     0.0000
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Overall_mean |
          y1 |   1.154606   .1860421     6.21   0.000     .7899701    1.519242
------------------------------------------------------------------------------

Estimated between-studies covariance matrix Sigma:
           y1
y1  .18579341
. mvmeta y S, vars(y2) bscov
Note: using method reml
Note: using variable y2
Note: 10 observations on 1 variables
  (output omitted)
                                                Number of obs   =         10
                                                Wald chi2(1)    =      12.93
Log likelihood = -18.728644                     Prob > chi2     =     0.0003
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Overall_mean |
          y2 |   1.963801   .5460555     3.60   0.000     .8935515     3.03405
------------------------------------------------------------------------------

Estimated between-studies covariance matrix Sigma:
           y2
y2  2.386426
```

These results agree with SAS PROC MIXED as reported by Riley et al. (2007b), except that the standard errors for the overall means are slightly larger (0.5461 for `y2`, compared with 0.5414 from SAS). This is because SAS does not, by default, allow for uncertainty in the estimated between-studies variance (SAS Institute 1999). `mvmeta`'s `nouncertainv` option inverts just the elements of the information matrix relating to the overall mean and agrees with SAS PROC MIXED:

```
. mvmeta y S, vars(y2) nouncertainv
Note: using method reml
Note: using variable y2
Note: 10 observations on 1 variables
  (output omitted)
Alternative standard errors, ignoring uncertainty in V:
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Overall_mean						
y2	1.963801	.5413727	3.63	0.000	.9027297	3.024872

5.2 Multivariate analysis

Because sensitivity and specificity are estimated on separate groups of individuals, their within-study covariance is zero. We could generate a new variable, `S12=0`, but it is easier to use the `corr(0)` option:

```
. mvmeta y S, corr(0) bscov
Note: using method reml
Note: using variables y1 y2
Note: 10 observations on 2 variables
Note: corr(0) used for all covariances
  (output omitted)
                                                Number of obs   =         10
                                                Wald chi2(2)    =     159.58
Log likelihood = -24.415968                     Prob > chi2     =     0.0000
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Overall_mean						
y1	1.166187	.1863275	6.26	0.000	.8009913	1.531382
y2	2.057752	.5607259	3.67	0.000	.9587493	3.156755

```
Estimated between-studies covariance matrix Sigma:
            y1         y2
y1  .20219111
y2  -.7227506  2.5835381
```

Again these results agree with those of Riley et al. (2007b), except that our standard errors are slightly larger because they allow for uncertainty in the between-studies covariance, Σ.

6 Example 2: Fibrinogen Studies Collaboration data

Fibrinogen Studies Collaboration (FSC) is a meta-analysis of individual data on 154,012 adults from 31 prospective studies with information on plasma fibrinogen and major disease outcomes (Fibrinogen Studies Collaboration 2004). As part of the published analysis, the incidence of coronary heart disease was compared across 10 groups defined

by baseline levels of fibrinogen (Fibrinogen Studies Collaboration 2005). That analysis used a fixed-effects model; here we allow for heterogeneity between studies by using a random-effects model, but we reduce the analysis to five groups to avoid presenting lengthy output.

In the first stage of analysis, we start with individual-level data including fibrinogen concentration, `fg`, in five levels. Following standard practice in the analysis of these data (Fibrinogen Studies Collaboration 2005), all analyses are stratified by sex and, for two studies that were randomized trials, by trial arm (variable `tr`). We adjust all analyses for age (variable `ages`), although in practice, more confounders would be adjusted for. We use the `esave(N)` option to record the sample size used in each study in variable `_e_N`.

```
. stset duration allchd
```

(*output omitted*)

```
. xi: mvmeta_make stcox ages i.fg, strata(sex tr) nohr
> saving(FSCstage1) replace by(cohort) usevars(i.fg) names(b V) esave(N)
i.fg              _Ifg_1-5            (naturally coded; _Ifg_1 omitted)
Using coefficients: _Ifg_2 _Ifg_3 _Ifg_4 _Ifg_5
-> cohort==1
         failure _d:  allchd
   analysis time _t:  duration
Iteration 0:   log likelihood = -5223.9564
Iteration 1:   log likelihood = -5135.3888
Iteration 2:   log likelihood = -5129.5633
Iteration 3:   log likelihood =  -5129.551
Refining estimates:
Iteration 0:   log likelihood =  -5129.551
Stratified Cox regr. -- Breslow method for ties
No. of subjects =        14436                     Number of obs   =     14436
No. of failures =          603
Time at risk    =  127969.6428
                                                   LR chi2(5)      =    188.81
Log likelihood  =    -5129.551                     Prob > chi2     =    0.0000
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ages	.0501925	.0072871	6.89	0.000	.03591	.064475
_Ifg_2	.2523666	.1895222	1.33	0.183	-.11909	.6238233
_Ifg_3	.5317069	.1804709	2.95	0.003	.1779905	.8854233
_Ifg_4	.9464425	.1761563	5.37	0.000	.6011824	1.291703
_Ifg_5	1.400935	.1779354	7.87	0.000	1.052188	1.749682

Stratified by sex tr

```
-> cohort==2
```

(*output omitted*)

Here are the data stored for the first 15 of the 31 studies; the data also include covariances `V_Ifg_2_Ifg_3`, etc., which are not displayed to save space. The first row of the data below reproduces the results from the `stcox` analysis given above.

```
. use FSCstage1, clear
. format b* V* %5.3f
. list cohort b_Ifg_2 b_Ifg_3 b_Ifg_4 b_Ifg_5 V_Ifg_2_Ifg_2 V_Ifg_3_Ifg_3,
> clean noobs
```

cohort	b_Ifg_2	b_Ifg_3	b_Ifg_4	b_Ifg_5	V_Ifg_~2	~3_Ifg_3
1	0.252	0.532	0.946	1.401	0.036	0.033
2	-0.184	-0.032	0.119	0.567	0.348	0.344
3	0.001	-0.529	-0.339	0.416	0.375	0.323
4	0.066	0.184	0.407	0.645	0.058	0.053
5	0.078	0.406	0.544	1.088	0.101	0.083
6	-0.113	0.456	0.456	0.875	0.065	0.054
7	-2.149	-0.264	-0.494	0.169	1.336	0.421
8	-0.039	0.170	0.420	1.053	0.042	0.038
9	0.443	0.595	0.922	0.797	0.202	0.175
10	0.356	1.312	0.628	2.133	1.500	1.170
11	1.297	1.052	1.421	1.752	0.559	0.542
12	0.323	0.545	0.681	0.540	0.132	0.122
13	-0.042	0.509	0.560	0.998	0.088	0.072
14	-2.667	-2.524	-2.010	-1.767	1.337	0.584
15	5.946	5.420	6.088	7.057	189.088	189.271

(*output omitted*)

Note the large parameter estimates and very large variances in study 15, which occur because this study has no events in category 1 of `fg`. Details of how such *perfect prediction* is handled are described in section 7.

Now the second stage of analysis:

```
. mvmeta b V
Note: using method reml
Note: using variables b_Ifg_2 b_Ifg_3 b_Ifg_4 b_Ifg_5
Note: 31 observations on 4 variables
```

(*output omitted*)

```
                                              Wald chi2(4)    =     139.59
Log likelihood = -79.489126                   Prob > chi2     =     0.0000
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Overall_mean						
b_Ifg_2	.1615842	.0796996	2.03	0.043	.005376	.3177925
b_Ifg_3	.3926019	.0878114	4.47	0.000	.2204947	.5647091
b_Ifg_4	.5620076	.0905924	6.20	0.000	.3844497	.7395654
b_Ifg_5	.8973289	.0942603	9.52	0.000	.712582	1.082076

```
Estimated between-studies SDs and correlation matrix:
               SD    b_Ifg_2    b_Ifg_3    b_Ifg_4    b_Ifg_5
b_Ifg_2  .22734097          1  .98953788  .97421937  .70621223
b_Ifg_3  .28611302  .98953788          1  .99657543  .80096928
b_Ifg_4  .30834247  .97421937  .99657543          1  .84773246
b_Ifg_5  .32742861  .70621223  .80096928  .84773246          1
```

It is interesting to compare the estimates with those obtained from four univariate meta-analyses, which can be run by `mvmeta b V, vars(b_Ifg_2)`, etc., and are summarized in table 1.

Table 1. Summary of estimates from four univariate meta-analyses

Group	Univariate			Multivariate						
	$\widehat{\mu}_i$	se($\widehat{\mu}_i$)	$\widehat{\tau}_i$	$\widehat{\mu}_i$	se($\widehat{\mu}_i$)	$\widehat{\tau}_i$	Correlations $\widehat{\kappa}_{ij}$			
2 vs 1	0.200	0.066	0.134	0.162	0.080	0.227	1			
3 vs 1	0.430	0.073	0.196	0.393	0.088	0.286	0.990	1		
4 vs 1	0.568	0.084	0.263	0.562	0.091	0.308	0.974	0.997	1	
5 vs 1	0.840	0.101	0.363	0.897	0.094	0.327	0.706	0.801	0.848	1

The univariate and multivariate methods give broadly similar point estimates, $\widehat{\mu}_i$, but the multivariate method gives rather larger estimates of three between-studies standard deviations, $\widehat{\tau}_i$, and, consequently, larger standard errors for $\widehat{\mu}_i$. A different choice of reference category would yield the same multivariate results but different univariate results. Of course, the multivariate method also has the advantage of estimating the between-studies correlations.

7 Perfect prediction

7.1 The problem

One difficulty that can occur in regression models with a categorical or time-to-event outcome is *perfect prediction* or *separation* (Heinze and Schemper 2002). In logistic regression, for example, perfect prediction occurs if there is a level of a categorical explanatory variable for which the observed values of the outcome are all one (or all zero); in Cox regression, it occurs if there is a category in which no events are observed. Here, as one or more regression parameters go to plus or minus infinity, the log likelihood increases to a limit and the second derivative of the log likelihood tends to zero.

Stata handles this problem in two ways. Stata first attempts to detect perfect prediction. If successful, it drops the relevant observations and term from the model. However, sometimes (in particular, if perfect prediction is in the reference category of a variable with more than two levels) Stata fails to detect perfect prediction. Here Stata reports very large ML estimates, observes that the variance–covariance matrix is singular, and reports a generalized inverse.

In the meta-analysis context, perfect prediction is likely to occur in some studies and not in others. (In the FSC analysis, it occurred in four studies.) Unfortunately, neither of the above solutions is satisfactory. In the first case, the model fit to a study with perfect prediction differs from that fit to other studies and has fewer parameters, so combination across studies is not meaningful. In the second case, some extremely large coefficients have inappropriately moderate standard errors, so they can have an excessive influence on meta-analytic results.

As an example, we use data from FSC study 15, which has no events in the reference category `fg==1`:

```
. xi: stcox ages i.fg if cohort==15, nohr
  (output omitted)
No. of subjects =         3134                     Number of obs   =      3134
No. of failures =           17
Time at risk    =  9465.954814
                                                   LR chi2(5)      =     16.43
Log likelihood  =   -127.22742                     Prob > chi2     =    0.0057
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ages	.0357279	.0263705	1.35	0.175	-.0159573	.087413
_Ifg_2	21.36403	.9147602	23.35	0.000	19.57113	23.15692
_Ifg_3	20.84916	.	.	.	.	.
_Ifg_4	21.50048	.8689028	24.74	0.000	19.79746	23.2035
_Ifg_5	22.47926	.7987255	28.14	0.000	20.91379	24.04473

Perfect prediction has not been detected, and the coefficients are appropriately large but with inappropriately small standard errors.

7.2 Solution: Augmentation

`mvmeta_make` checks for perfect prediction by checking that 1) all parameters are reported and 2) there are no zeros on the diagonal of the variance–covariance matrix of the parameter estimates. If perfect prediction is detected, `mvmeta_make` augments the data in such a way as to avoid perfect prediction but gives the added observations a tiny weight to minimize their impact on well-estimated parts of the model.

The augmentation is performed at two design points for each covariate x, defined by letting $x = \overline{x} \pm s_x$ (where $\overline{x}$ and s_x are the study-specific mean and standard deviation of x, respectively) and by fixing other covariates at their mean value. The records added at each design point depend on the form of regression model. For logistic regression, we add one event and one nonevent. For other regression models with discrete outcomes, we add one observation with each outcome level. For survival analyses, we add one event at time $t_{\min}/2$ and one censoring at time $t_{\max} + t_{\min}/2$, where $t_{\min}$ and $t_{\max}$ are the first and last follow-up times in the study. For a stratified Cox model, the augmentation is performed for each stratum.

A total weight of wp is then shared equally between the added observations, where w is specified by the `augwt()` option (the default is `augwt(0.01)`), and p is the number of model parameters (treating the baseline hazard in a Cox model as one parameter). The regression model is then rerun including the weighted added observations. For study 15, this yields

(*Continued on next page*)

```
No. of subjects =      3134.06                 Number of obs   =      3134
No. of failures =        17.03
Time at risk    =  9466.077771
                                               LR chi2(5)      =     16.33
Log likelihood  =   -115.75111                 Prob > chi2     =    0.0060
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ages	.0353976	.0263231	1.34	0.179	-.0161948	.08699
_Ifg_2	5.946375	13.75093	0.43	0.665	-21.00495	32.89771
_Ifg_3	5.41975	13.75757	0.39	0.694	-21.54459	32.38409
_Ifg_4	6.088434	13.74965	0.44	0.658	-20.86039	33.03726
_Ifg_5	7.057288	13.74605	0.51	0.608	-19.88448	33.99905

```
                                                Stratified by sex tr
```

The coefficients for the `_Ifg_*` terms are reduced but still large, but their large standard errors now mean that they will not unduly influence the meta-analysis. The coefficient and standard error for `ages` are barely changed. It is useful to compare the variance–covariance matrix of the parameter estimates before augmentation,

```
ages      _Ifg_2      _Ifg_3      _Ifg_4      _Ifg_5
  ages   .00069444
_Ifg_2   .00156723   .83711768
_Ifg_3           0           0           0
_Ifg_4  -.00185585   .49628548           0   .75596628
_Ifg_5  -.00303957   .49370111           0   .50944939   .64022023
```

with that after augmentation:

```
ages      _Ifg_2      _Ifg_3      _Ifg_4      _Ifg_5
  ages   .00069291
_Ifg_2  -.00309014   189.08811
_Ifg_3  -.00465418   188.76205   189.27067
_Ifg_4  -.00650648   188.77085   188.78488   189.05294
_Ifg_5  -.00768805   188.77649   188.79309   188.81504   188.95394
```

Because the covariances in the latter matrix are large, contrasts between groups 2, 3, 4, and 5 will receive appropriately small standard errors. This study will therefore contribute information about contrasts between groups 2, 3, 4, and 5 to the meta-analysis, but it will contribute no information about contrasts between group 1 and other groups.

A related problem occurs if some study has no observations at all in a particular category. The augmentation algorithm is applied here, too, with the modification that the value s_x, used to define the added design points, is taken as the standard deviation across all studies, because the within-study standard deviation is zero.

8 Discussion

8.1 Difficulties and limitations

The main difficulty that might be encountered in fitting multivariate random-effects meta-analysis models is a nonpositive-definite Σ. However, the parameterization used here ensures that Σ is positive semidefinite and achieves a nonpositive-definite Σ if one or more elements of the Cholesky decomposition approach zero. I have encountered nonconvergence of the Newton–Raphson algorithm only when the starting value is $\Sigma = 0$, which is avoided by a suitable nonzero choice of starting values, or when inappropriately handled perfect prediction has led to extreme parameter estimates with small standard errors.

The standard error provided for an REML analysis allows for uncertainty in estimating Σ by inverting the second derivative matrix of the restricted likelihood (1). This is not the standard approach (Kenward and Roger 1997), and its properties require further investigation. Confidence intervals based on a t distribution would be a useful enhancement.

At present, the augmentation routine in `mvmeta_make` effectively ignores any category in which perfect prediction occurs but allows information to be drawn from other categories from that study. A larger augmentation would allow information to be drawn from categories with perfect prediction. For example, if the data consist of 2×2 tables, then standard practice would add 0.5 observations to each cell (Sweeting, Sutton, and Lambert 2004). This amounts to assigning to the augmented observations a total weight equal to the number of parameters, and it is tempting to apply this rule more widely (by using `augment(1)`). However, larger augmentation weights have the undesirable property of not being invariant to reparameterization; for example, a different choice of reference category for the `fg` variable in section 6 would lead to somewhat different results. Larger augmentation is probably best implemented by the user.

There are alternate ways to handle perfect prediction, including various forms of penalized likelihood. The methods of Le Cessie and van Houwelingen (1992) and Verweij and van Houwelingen (1994) have been implemented in Stata by the `plogit` and `stpcox` commands, respectively, and both are currently being updated to allow for perfect prediction (G. Ambler, pers. comm.). The method of Firth (1993) is invariant to reparameterization and is being implemented by the author. When suitable routines become available in Stata, they can be called by the `ppcmd()` option in `mvmeta_make`.

8.2 Comparison to other procedures

All the models considered here can also be fit in SAS PROC MIXED, although some programming effort is required to specify the known within-study variances, S_i. The two approaches are very similar, but by default, SAS produces standard errors that ignore the uncertainty in Σ, and produces confidence intervals by using the t distribution on

$n-1$ degrees of freedom. Further, SAS optionally provides a standard error adjusted to allow for uncertainty in estimating Σ and provides the approximate degrees of freedom of Kenward and Roger (1997), which has good small-sample properties.

Multivariate meta-analysis models cannot be fit by using existing Stata commands, but univariate models can. `metan` differs from `mvmeta` because it uses DerSimonian and Laird (1986) estimation of the random-effects variance. `metareg` offers the choice of DerSimonian and Laird, ML, or REML estimation, so if run without covariates, it can be compared to `mvmeta`. The original `metareg` (Sharp 1998) used the algorithm of Hardy and Thompson (1996) and did not always find the best solution. Version 2 of `metareg`, by Harbord and Higgins (2008), uses Newton–Raphson maximization via `ml`, and produces the same point estimates as `mvmeta` and the same standard errors as `mvmeta` with the `nouncertainv` option. `metareg` produces confidence intervals that allow for nonnormality of the sampling distributions by using the method of Knapp and Hartung (2003); its `z` option produces confidence intervals that agree with `mvmeta`. Of course, `metareg` also has the enormous advantage of handling meta-regression.

8.3 More than two outcomes

Although `mvmeta` handles several outcomes perfectly well, its computing time increases sharply as the number of outcomes increases. `mvmeta` can even computationally handle situations where there are more quantities of interest than studies ($p > n$); however, fitting such large models can be unwise and results can be untrustworthy.

9 Acknowledgments

I thank the FSC for providing access to their data for illustrative analyses: a full list of the FSC collaborators is given in Fibrinogen Studies Collaboration (2005). I also thank Li Su and Dan Jackson for helping me rediscover the Cholesky decomposition parameterization; Stephen Kaptoge and Sebhat Erquo for helpful comments on the programming; James Roger for help in understanding SAS PROC MIXED; and Patrick Royston and Gareth Ambler for discussions about augmentation and penalized likelihoods.

10 References

Berkey, C. S., J. J. Anderson, and D. C. Hoaglin. 1996. Multiple-outcome meta-analysis of clinical trials. *Statistics in Medicine* 15: 537–557.

Berkey, C. S., D. C. Hoaglin, A. Antczak-Bouckoms, F. Mosteller, and G. A. Colditz. 1998. Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine* 17: 2537–2550.

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

Fibrinogen Studies Collaboration. 2004. Collaborative meta-analysis of prospective studies of plasma fibrinogen and cardiovascular disease. *European Journal of Cardiovascular Prevention and Rehabilitation* 11: 9–17.

———. 2005. Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: An individual participant meta-analysis. *Journal of the American Medical Association* 294: 1799–1809.

Firth, D. 1993. Bias reduction of maximum likelihood estimates. *Biometrika* 80: 27–38.

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* 30: 619–629.

Hasselblad, V. 1998. Meta-analysis of multitreatment studies. *Medical Decision Making* 18: 37–43.

Heinze, G., and M. Schemper. 2002. A solution to the problem of separation in logistic regression. *Statistics in Medicine* 21: 2409–2419.

Jackson, D., I. R. White, and S. G. Thompson. Forthcoming. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Statistics in Medicine*.

Kenward, M. G., and J. H. Roger. 1997. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* 53: 983–997.

Knapp, G., and J. Hartung. 2003. Improved tests for a random-effects meta-regression with a single covariate. *Statistics in Medicine* 22: 2693–2710.

Le Cessie, S., and J. C. van Houwelingen. 1992. Ridge estimators in logistic regression. *Applied Statistics* 41: 191–201.

Lu, G., and A. E. Ades. 2004. Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine* 23: 3105–3124.

Normand, S. L. T. 1999. Meta-analysis: Formulating, evaluating, combining and reporting. *Statistics in Medicine* 18: 213–259.

Pinto, E., A. Willan, and B. O'Brien. 2005. Cost-effectiveness analysis for multinational clinical trials. *Statistics in Medicine* 24: 1965–1982.

Riley, R. D., K. R. Abrams, P. C. Lambert, A. J. Sutton, and J. R. Thompson. 2007a. An evaluation of bivariate random-effects meta-analysis for the joint synthesis of two correlated outcomes. *Statistics in Medicine* 26: 78–97.

Riley, R. D., K. R. Abrams, A. J. Sutton, P. C. Lambert, and J. R. Thompson. 2007b. Bivariate random-effects meta-analysis and the estimation of between-study correlation. *BMC Medical Research Methodology* 7: 3.

SAS Institute. 1999. SAS *OnlineDoc Version Eight*. Cary, NC: SAS Institute. http://www.technion.ac.il/docs/sas/.

Sharp, S. 1998. sbe23: Meta-analysis regression. *Stata Technical Bulletin* 42: 16–22. Reprinted in *Stata Technical Bulletin Reprints*, vol. 7, pp. 148–155. College Station, TX: Stata Press.

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.

Thompson, S. G. 1994. Why sources of heterogeneity in meta-analysis should be investigated. *British Medical Journal* 309: 1351–1355.

van Houwelingen, H. C., L. R. Arends, and T. Stijnen. 2003. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* 21: 589–624.

Verweij, P. J. M., and J. C. van Houwelingen. 1994. Penalized likelihood in Cox regression. *Statistics in Medicine* 13: 2427–2436.

About the author

Ian White is a senior statistician at the MRC Biostatistics Unit in Cambridge, UK. His research interests include missing data, noncompliance and measurement error in clinical trials, observational studies, and meta-analysis.