OPEN ACCESS

University of BRISTOL

Chaimani, A., Mavridis, D., Higgins, J., Salanti, G., & White, I. R. (2018). Allowing for informative missingness in aggregate data meta-analysis with continuous or binary outcomes: Extensions to metamiss. *Stata Journal*, *18*(3), 716-740. [st0540].

Peer reviewed version

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

Meta-analysis with missing continuous or binary data: extensions to metamiss

Anna Chaimani, Dimitris Mavridis, Julian Higgins, Georgia Salanti, Ian R. White

Anna Chaimani*
Paris Descartes University;
INSERM, UMR1153 Epidemiology and
Statistics, Sorbonne Paris Cité Research
Center (CRESS), METHODS Team;
Cochrane France
Paris, France
anna.chaimani@parisdescartes.fr

Dimitris Mavridis
Department of Primary Education,
School of Education
University of Ioannina
Ioannina, Greece
dmavridi@cc.uoi.gr

Julian P. T. Higgins
Population Health Sciences,
Bristol Medical School
University of Bristol
Bristol, UK
julian.higgins@bristol.ac.uk

Georgia Salanti
Institute of Social and Preventive Medicine
University of Bern
Bern, Switzerland
georgia.salanti@ispm.unibe.ch

Ian R. White
MRC Biostatistics Unit
Cambridge, UK
and
MRC Clinical Trials Unit at UCL
London, UK
ian.white@ucl.ac.uk

*Corresponding author

Abstract

Missing outcome data can invalidate the results of randomized trials and their meta-analysis. The impact of missing outcome data on the summary effect can be explored by assuming a relationship between the outcome in the observed and the missing participants via an informative missingness parameter (IMP). The use of IMPs in pairwise meta-analysis has been previously implemented in Stata with the `metamiss` command for the case of binary outcome data. This paper presents the new command `metamiss2`, which is an extension of `metamiss` for binary or continuous data in pairwise or network meta-analysis.

1. Introduction

Missing outcome data are a common threat to the validity of randomized trials and, subsequently, their meta-analysis. Because missing data are by definition not present in the data set, addressing them requires making untestable assumptions. Researchers typically

ignore missing data and analyse complete data only, which is equivalent to assuming that missing participants are missing at random (MAR) (Little and Rubin, 2014); we refer to such an analysis as an available cases analysis (ACA). However, this approach can give biased results when missingness is associated with the actual outcome. This is the case of data missing not at random (MNAR).

Here we consider randomized trials with an outcome measured at a single time point, for which outcome data are unavailable for some of the participants within the trial. Furthermore, we focus on approaches that are based on aggregated (summary) data from the trial, such as are often available from journal articles, and which are typical of the data available for a meta-analysis.

The use of pattern mixture models has been previously suggested for handling missing outcome data in meta-analysis of binary outcomes with aggregate data (White et al., 2008a). This approach is based on the *informative missingness odds ratio* (IMOR), which relates the odds of outcome in the missing data to that in the observed data. The approach can allow for uncertainty in the IMOR, and has been implemented in Stata in the `metamiss` command.

Parameters like the IMOR that measure departure from a MAR assumption have been called "sensitivity parameters" by Kenward et al (2001); we follow White et al (2011) in calling them "informative missingness parameters" (IMPs). Mavridis et al. extended the IMP framework to meta-analyses with continuous outcomes by defining IMPs that relate the mean of the outcome between the missing and the observed participants (Mavridis et al., 2015).

Network meta-analysis (NMA) combines the results of multiple direct comparisons (Salanti et al., 2008) and is therefore prone to the same biases as pairwise meta-analysis. More specifically, incorrectly handling missing data in one or more comparisons of a NMA could affect all relative effects in which this particular comparison is involved either directly or indirectly (Salanti et al., 2014). Methods used to allow for IMPs in pairwise MA apply directly to NMA when only two-group trials are included. In the presence of multi-group trials the 'adjusted' covariance between relative effects from the same study also needs to be estimated (Mavridis et al., 2015). The application of IMPs in NMA with binary outcome data has been exemplified in a Bayesian framework by Spineli et al (Spineli et al., 2013).

The aim of this paper is to introduce a new Stata command, called `metamiss2`, with new syntax, which extends `metamiss` (1) by handling continuous as well as binary outcome data and (2) by working in NMA as well as pairwise meta-analysis. `metamiss2` performs a two-stage approach: stage 1 estimates the 'adjusted' study-specific relative effects and their variances and covariances, and stage 2 calls `metan` (Harris et al., 2008) (for pairwise meta-analysis) or `network meta` (White, 2015) (for NMA) to obtain the summary effects.

2. Theory

This section describes stage 1 of the analysis, which estimates the treatment effect(s) and their variance(s) for each study allowing for MNAR data. The second stage combines the first-stage estimates: this is a standard meta-analysis procedure (Palmer and Sterne, 2016) and is not further described here. We describe first the case of two-group studies and then the extension for multi-group studies.

2.1. Two-group studies

Our notation follows that of (Mavridis et al., 2015), but is extended to cover the case of binary data as in (White et al., 2008a). We assume we have data from multiple studies, each with two groups denoted T (treatment) and C (control). We assume that in the jth group of the ith study ($j = C, T$), we know n_{ij}, the number of participants providing outcome data, and m_{ij}, the number of participants with missing outcome data. If the outcome is binary we assume we know r_{ij}, the number of observed successes. If the outcome is continuous we assume we know x_{ij}^{obs}, the mean of the observed outcomes, and s_{ij}, the standard deviation (SD) of the observed outcomes.

The model for the observed data is $r_{ij} \sim Bin(n_{ij}, \chi_{ij}^{obs})$ if the outcome is binary and $x_{ij}^{obs} \sim N(\chi_{ij}^{obs}, s_{ij}^2)$ if the outcome is continuous. In both cases χ_{ij}^{obs} is the true mean of the observed outcomes in the jth group of the ith study.

Our measure of interest in the ith study is defined as $\beta_i = f(\chi_{iT}^{tot}) - f(\chi_{iC}^{tot})$, where χ_{ij}^{tot} is the true mean outcome of all (observed and missing) outcomes in the jth group of the ith study. For a binary outcome, the link function f may be the identity function $f(x) = x$ (so that the measure of interest is the risk difference), the logarithmic function $f(x) = \log(x)$ (giving the log risk ratio) or the logit function $f(x) = logit(x)$ (giving the log odds ratio). For a continuous outcome, f may be the identity function $f(x) = x$ (giving the mean difference (MD)) or the logarithmic function $f(x) = \log(x)$ (giving the log ratio of means (RoM)); alternatively, $f(\chi)$ may be replaced by $f_i(\chi) = \chi/\sigma_i$ where σ_i is the pooled SD in the ith study, giving the standardised mean difference (SMD) (White and Thomas, 2005).

In this simple setting, a MAR assumption would state that $\chi_{ij}^{tot} = \chi_{ij}^{obs}$, which would justify ignoring the missing data (Little and Rubin, 2014). Under MNAR, we cannot ignore the missing data. Instead we view the mean outcome of all participants as a mixture of outcomes in the observed and in the missing participants. We write $\chi_{ij}^{tot} = \pi_{ij}\chi_{ij}^{obs} + (1 - \pi_{ij})\chi_{ij}^{mis}$ where π_{ij} is the probability of observed data and χ_{ij}^{mis} is the (unobserved) mean outcome in the missing data. We introduce the IMP as $\lambda_{ij} = g(\chi_{ij}^{mis}) - g(\chi_{ij}^{obs})$. For a binary outcome, we consider the case where g is the logit function $g(x) = logit(x)$ and the IMP is the log of IMOR (White et al., 2008a). For a continuous outcome, g may be the identity function $g(x) = x$ (so the IMP is the informative missingness difference of means, IMDoM), or the logarithmic function $g(x) = \log(x)$ (so the IMP is the log of the informative missingness ratio of means, log IMRoM) (Mavridis et al., 2015). We generally expect researchers to employ IMDoM with MD and SMD and IMRoM with RoM. When $\lambda_{ij} = 0$, we assume that the outcome in the missing participants is on average the same with the outcome in the observed outcomes. This is equivalent to assuming that we can safely ignore the missing data, which is the MAR assumption. We quantify departures from the MAR assumption by allowing λ_{ij} to assume non-zero values.

The IMPs λ_{ij} are required to estimate β_i but are not identified by the observed data. Instead, they are either specified by the analyst on the basis of subject-matter knowledge or a range of values is assumed in a sensitivity analysis. By allowing for uncertainty in the IMPs, the model reduces the relative weight given to studies with more missing data (White et al., 2008a). The IMPs may be independent across groups, $\lambda_{ij} \sim N\left(\mu_{\lambda_{ij}}, \sigma_{\lambda_{ij}}^2\right)$, or we can allow for correlation by assuming a bivariate normal distribution with $corr(\lambda_{iT}, \lambda_{iC}) = \rho_{\lambda_i}$. Thus, non-zero values of any of $\mu_{\lambda_{iT}}, \mu_{\lambda_{iC}}, \sigma_{\lambda_{iT}}^2, \sigma_{\lambda_{iC}}^2$ imply MNAR. IMPs are assumed to be

independent across studies to abide by the fundamental assumption of independent studies in meta-analysis.

Two estimation procedures are described briefly here and in more detail in Mavridis et al (2015). We write $\beta_i = \beta_i(\theta_i)$ where $\theta_i = \left(\pi_{iT}, \pi_{iC}, \chi_{iT}^{obs}, \chi_{iC}^{obs}, \lambda_{iT}, \lambda_{iC}\right)$. In the Taylor method, which uses a linear approximation to $\beta_i(\theta_i)$, the point estimate of β_i is $\hat{\beta}_i = \beta_i(\hat{\theta}_i)$ where $\hat{\theta}_i = \left(\hat{\pi}_{iT}, \hat{\pi}_{iC}, \hat{\chi}_{iT}^{obs}, \hat{\chi}_{iC}^{obs}, \mu_{\lambda_{iT}}, \mu_{\lambda_{iC}}\right)$, and its estimated variance is $\widehat{var}\left(\hat{\beta}_i\right) = D_i^T V_i D_i$ where $D_i = \frac{d\beta_i}{d\theta_i}$ evaluated at $\theta_i = \hat{\theta}_i$ and $V_i = \widehat{var}(\hat{\theta}_i)$ relates to the sampling variance for $\hat{\pi}_{iT}, \hat{\pi}_{iC}, \hat{\chi}_{iT}^{obs}, \hat{\chi}_{iC}^{obs}$ and the uncertainty variance for $\mu_{\lambda_{iT}}, \mu_{\lambda_{iC}}$. In the parametric bootstrap method, which avoids the linear approximation to $\beta_i(\theta_i)$, values θ_i^* are repeatedly drawn - $\pi_{iT}, \pi_{iC}, \chi_{iT}^{obs}, \chi_{iC}^{obs}$ independently from their posterior distributions given the data, and $\lambda_{iT}, \lambda_{iC}$ jointly from their prior distribution - and the point estimate $\hat{\beta}$ and its estimated variance are the mean and variance of the $\beta_i(\theta_i^*)$. When the measure of interest is the SMD, the procedure takes σ_i as the pooled SD across groups and ignores uncertainty in σ_i.

2.2. Multi-group studies

The same methods are used for multi-group studies, which may arise in network meta-analysis. Multi-group studies yield multiple treatment effects, e.g. $\beta_{i1} = f(\chi_{iT1}^{tot}) - f(\chi_{iC}^{tot})$ and $\beta_{i2} = f(\chi_{iT2}^{tot}) - f(\chi_{iC}^{tot})$. Extending the estimation method above yields estimates of their variances and the covariance $cov(\hat{\beta}_{i1}, \hat{\beta}_{i2})$ (Mavridis et al., 2015).

3. Syntax

```
metamiss2 [varlist] [if exp] [in range] [, imptype(imdom|logimrom)
   impmean(# #...#) impsd(# #...#) impcorrelation(real|exp|matrix)
   fixed inconsistency md smd rom or rr rd sdpool(on|off) taylor
   bootstrap reps(integer) compare(string) sensitivity nokeep nometa
   varchange netplot trtlabels(string) metanoptions(meta_options)
   networkoptions(network_meta_options) netplotreference(string)
   netplotoptions(intervalplot_options)]
```

where `varlist` is:

- for pairwise meta-analysis with continuous outcome data: `nE mE yE sdE nC mC yC sdC` - variables containing the numbers of observed and missing participants, the mean and SD of the observed data in experimental and control group respectively.

- for pairwise meta-analysis with dichotomous outcome data: `rE fE mE rC fC mC` - variables containing the numbers of successes and failures in the observed data and the number of missing participants in experimental and control group respectively.

- for network meta-analysis: `varlist` is not used but the data must have been prepared using the `network setup` command (White, 2015) in the augmented format (see example 4.3).

Options for specifying the IMPs

`imptype(imdom|logimrom)` specifies the type of IMP for continuous outcome data. The default is `imdom`. This option is not needed for dichotomous outcome data since the only available IMP is `logimor`.

`impmean(exp exp...exp)` specifies the mean of the assumed (normal) distribution for IMP. The default value is 0 in all groups. If one value is given, it is the mean for all groups. For pairwise meta-analysis, if two values are given, they are the means for the experimental and control group. For network meta-analysis, if T values are given (with T the total number of treatments), they are the means for the reference treatment and the non-reference treatments in the order shown in `network setup`. Each `exp` may be a single value corresponding to all studies or a variable containing study-specific values.

`impsd(exp exp...exp)` specifies the SD of the assumed (normal) distribution for IMP in the same way as described above for `impmean()`. The default value is 0 in all groups.

`impcorrelation(real|exp|matrix)` specifies the correlation of the IMP between the different groups. The default value is 0. A common correlation value for all pairs of treatments or the full $T \times T$ correlation matrix (only for network meta-analysis) can be specified.

`compare(string)` specifies a second assumption for IMP to be compared to the primary analysis. `string` may include `impmean()`, `impsd()` and `impcorrelation()`.

`sensitivity` specifies a sensitivity analysis for the IMP assuming a range of different standard deviations for its distribution with `impmean(0)` or a different specified `impmean()`.

Options for continuous data

`smd` specifies the standardized mean difference as the measure of interest (the default for continuous data).

`md` specifies the mean difference as the measure of interest.

`rom` specifies the ratio of means as the measure of interest.

`sdpool(on|off)` specifies whether the SD is pooled across groups in computing variances. Following `metan`, the default option for mean difference and ratio of means is `sdpool(off)` and for standardized mean difference is `sdpool(on)`.

Options for dichotomous data

`rr` specifies the risk ratio (RR) as the measure of interest (the default for dichotomous data). Note that in this case the IMP is the `logimor`.

`or` specifies the odds ratio as the measure of interest. Note that in this case the IMP is the `logimor`.

`rd` specifies the risk difference as the measure of interest. Note that in this case the IMP is `logimor`.

Estimation options

`taylor` specifies that Taylor-series approximation is used to integrate over the distribution of the IMP (the default).

`bootstrap` specifies that parametric bootstrap is used to integrate over the distribution of the IMP.

`reps(integer)` specifies the number of simulations under the bootstrap method. The default is 10000.

Meta-analysis options

`fixed` specifies the use of the fixed-effect model instead of the default random-effects model.

`inconsistency` specifies the use of an inconsistency model for the case of network meta-analysis instead of the consistency model which is the default.

`nometa` skips the conduct of pairwise or network meta-analysis after estimating the 'adjusted' study-specific effect sizes and variances.

`metanoptions(meta_options)` are any valid options for `metan`.

`networkoptions(network_meta_options)` are any valid options for `network meta`.

Output options

`nokeep` specifies that study-specific 'adjusted' effect sizes and standard errors/variances will be dropped from the dataset. By default, these estimates are stored as extra variables; for pairwise meta-analysis with names `_ES`, `_seES` (as in `metan`) and in network meta-analysis with prefix `_imp_`.

`varchange` specifies that the 'adjusted' study-specific relative effects and variances are stored in the dataset replacing the respective values obtained from the `network setup` command. This means that the current assumptions about the missing data will also apply to future analyses of the data.

`netplot` specifies that a forest plot with the relative effects from network meta-analysis will be drawn. The same forest plot can be produced by running the `intervalplot` command (Chaimani and Salanti, 2015) after `metamiss2` for a network meta-analysis. Note that for the case of pairwise meta-analysis a forest plot is produced by default.

`trtlabels(string)` specifies the labels of the treatments for the case of network meta-analysis separated with spaces. These labels will be used in `netplot`. The first label should correspond to the reference treatment and the other treatment should be given in the numerical/alphabetical order of their codes in the data.

`netplotreference(string)` specifies a treatment to be used as reference in the `netplot` so as only a subset of the relative effects from the network meta-analysis (i.e. every treatment vs. that reference) will be given in the forest plot. The treatment specified here can be different from the reference treatment of the analysis.

`netplotoptions(intervalplot_options)` specifies any additional valid option allowed in `intervalplot`.

4. Examples

4.1. Pairwise meta-analysis, binary data

We illustrate the use of `metamiss2` for pairwise meta-analysis with binary outcome data using a dataset that includes 17 trials comparing the effectiveness of haloperidol against placebo for the treatment of schizophrenia. The outcome is clinical response and RR>1 suggest that haloperidol works better that placebo.

```
. li,clean noo

       author   year   rh   fh   mh   rp   fp   mp
    Arvanitis   1997   25   25    2   18   33    0
```

```
          Beasley   1996   29   18   22   20   14   34
         Bechelli   1983   12   17    1    2   28    1
          Borison   1992    3    9    0    0   12    0
        Chouinard   1993   10   11    0    3   19    0
           Durost   1964   11    8    0    1   14    0
            Garry   1962    7   18    1    4   21    1
           Howard   1974    8    9    0    3   10    0
           Marder   1994   19   45    2   14   50    2
     Nishikawa_82   1982    1    9    0    0   10    0
     Nishikawa_84   1984   11   23    3    0   13    0
          Reschke   1974   20    9    0    2    9    0
           Selman   1976   17    1   11    7    4   18
    Serafetinides   1972    4   10    0    0   13    1
          Simpson   1967    2   14    0    0    7    1
          Spencer   1992   11    1    0    1   11    0
         Vichaiya   1971    9   20    1    0   29    1
```

We explore different assumptions about the association of the outcome between missing and observed data, which we describe by the $logIMOR$.

First, we assume that our beliefs about the missing data can be expressed as follows. In the haloperidol group we believe there may be systematic differences between outcomes in missing and observed participants but we are not sure in which direction, so we give the $logIMOR$ a distribution with mean 0 and SD 1. In the placebo group we believe the response in missing participants is probably worse than in observed participants, so we give the $logIMOR$ a distribution with mean -1 and SD 1. This can be the case when for example patients dropout the study because their symptoms have worsen. We use the default method of estimation, which is Taylor-series approximation. We use the metan options `lcols(author)` to label the studies.

```
. metamiss2 rh fh mh rp fp mp, impmean(0 -1) impsd(1) metanopt(lcols(author))

********************************************************************
******** METAMISS2: meta-analysis allowing for missing data *******
******** Informative missingness parameter with uncertainty *******
********************************************************************

Informative missingness parameter: logIMOR
Measure of interest:               Risk ratio
Assumed distribution for IMP:      Experimental group ~ N(0,1^2)
                                   Control group ~ N(-1,1^2)
IMP correlation between groups:    0
Method for first stage model:      Taylor series approximation
Second stage model:                Random effects meta-analysis

(Calling metan with options: lcols(author) ...)
```

Study	ES	[95% Conf. Interval]		% Weight
Arvanitis	1.417	0.890	2.256	18.58
Beasley	1.323	0.720	2.432	14.50
Bechelli	6.333	1.547	25.918	4.39
Borison	7.000	0.400	122.442	1.19
Chouinard	3.492	1.113	10.955	6.21
Durost	8.684	1.258	59.946	2.51
Garry	1.791	0.596	5.381	6.60
Howard	2.039	0.670	6.208	6.48
Marder	1.381	0.758	2.517	14.72
Nishikawa_82	3.000	0.137	65.903	1.03
Nishikawa_84	9.200	0.580	146.044	1.28
Reschke	3.793	1.058	13.604	5.19
Selman	1.949	0.906	4.194	11.09

```
Serafetinides        |  8.764      0.516   148.917         1.22
Simpson              |  2.526      0.135    47.152         1.14
Spencer              | 11.000      1.671    72.396         2.62
Vichaiya             | 19.393      1.180   318.749         1.25
---------------------+-------------------------------------------------
D+L pooled ES        |  2.211      1.607     3.042       100.00
---------------------+-------------------------------------------------

  Heterogeneity chi-squared =  20.66 (d.f. = 16) p = 0.192
  I-squared (variation in ES attributable to heterogeneity) =  22.6%
  Estimate of between-study variance Tau-squared =  0.0863

  Test of ES=1 : z=   4.87 p = 0.000
```

After running `metamiss2` the 'adjusted' study-specific relative effects along with their 95% confidence intervals are given in the output. The same results are obtained when we run the same analysis with `metamiss`:

```
. metamiss rh fh mh rp fp mp, logimor(0 -1) sdlogimor(1) method(Taylor) randomi
lcols(author)


*******************************************************************
******** METAMISS: meta-analysis allowing for missing data ********
********           Bayesian analysis using priors          ********
*******************************************************************
Measure: RR.
Zero cells detected: adding 1/2 to 6 studies.
Priors used:  Group 1: N(0,1^2). Group 2: N(-1,1^2). Correlation: 0.
Method: Taylor series approximation.

(Calling metan with options:  randomi lcols(author) eform  ...)

{output omitted}
```

The above analysis implicitly assumes that the IMPs in the two groups are unrelated. We next assume that a high $logIMOR$ in one group is likely to go with a high $logIMOR$ in the other group; that entails the two $logIMOR$s are positively correlated, with correlation $\rho = 0.5$. We obtain the study-specific RRs using the bootstrap method:

```
. metamiss2 rh fh mh rp fp mp, impmean(0 -1) impsd(1) impc(0.5) b metanopt(lcols(author))

*******************************************************************
******** METAMISS2: meta-analysis allowing for missing data *******
******** Informative missingness parameter with uncertainty *******
*******************************************************************

Informative missingness parameter: logIMOR
Measure of interest:                Risk ratio
Assumed distribution for IMP:       Experimental group ~ N(0,1^2)
                                    Control group ~ N(-1,1^2)
IMP correlation between groups:     0.5
Method for first stage model:       Parametric Bootstrap (10000 draws)
Second stage model:                 Random effects meta-analysis

(Calling metan with options: lcols(author) ...)

           Study     |    ES    [95% Conf. Interval]    % Weight
---------------------+-------------------------------------------------
Arvanitis            |  1.426      0.888     2.290        17.68
Beasley              |  1.298      0.760     2.217        16.40
Bechelli             |  7.898      1.600    38.975         4.52
Borison              | 21.875      0.258  1853.641         0.70
Chouinard            |  3.959      1.133    13.836         6.60
Durost               | 14.440      1.218   171.198         2.11
```

```
Garry              |  1.888     0.580     6.149           7.17
Howard             |  2.275     0.668     7.741           6.81
Marder             |  1.398     0.760     2.574          14.92
Nishikawa_82       |  7.395     0.073   744.881           0.65
Nishikawa_84       | 30.602     0.406  2306.082           0.73
Reschke            |  4.733     1.108    20.229           5.26
Selman             |  1.973     0.919     4.236          12.19
Serafetinides      | 26.937     0.348  2082.840           0.72
Simpson            |  6.696     0.077   579.813           0.69
Spencer            | 18.484     1.575   216.868           2.13
Vichaiya           | 67.066     0.852  5281.565           0.72
-------------------+------------------------------------------------
D+L pooled ES      |  2.342     1.609     3.411         100.00
-------------------+------------------------------------------------

  Heterogeneity chi-squared =  23.53 (d.f. = 16) p = 0.100
  I-squared (variation in ES attributable to heterogeneity) =  32.0%
  Estimate of between-study variance Tau-squared =  0.1496

  Test of ES=1 : z=   4.44 p = 0.000
```

Running the same analysis with `metamiss` gives slightly different results:

```
. metamiss rh fh mh rp fp mp, logimor(0 -1) sdlogimor(1) corrlogimor(0.5) method(mc)
randomi lcols(author) reps(10000)
********************************************************************
******** METAMISS: meta-analysis allowing for missing data ********
********          Bayesian analysis using priors          ********
********************************************************************
Measure: logRR.
Zero cells detected: adding 1/2 to 6 studies.
Priors used:  Group 1: N(0,1^2). Group 2: N(-1,1^2). Correlation: 0.5.
Method: Monte Carlo (10000 draws).
..................................................................................
>.................................................................................
{output omitted}

(Calling metan with options:  randomi lcols(author) eform  ...)

          Study    |    ES    [95% Conf. Interval]    % Weight
-------------------+------------------------------------------------
Arvanitis          |  1.420     0.891     2.263          17.10
Beasley            |  1.302     0.763     2.221          14.92
Bechelli           |  5.322     1.504    18.828           4.32
Borison            |  4.025     0.523    30.955           1.81
Chouinard          |  3.173     1.106     9.105           5.87
Durost             |  6.204     1.325    29.046           3.03
Garry              |  1.736     0.619     4.864           6.09
Howard             |  1.948     0.684     5.551           5.93
Marder             |  1.374     0.762     2.477          13.38
Nishikawa_82       |  2.046     0.225    18.563           1.56
Nishikawa_84       |  5.125     0.757    34.711           2.04
Reschke            |  3.495     1.113    10.975           5.12
Selman             |  1.953     0.924     4.129           9.85
Serafetinides      |  4.822     0.676    34.407           1.94
Simpson            |  1.661     0.214    12.870           1.80
Spencer            |  7.757     1.740    34.584           3.21
Vichaiya           |  9.990     1.477    67.586           2.04
-------------------+------------------------------------------------
D+L pooled ES      |  2.143     1.616     2.842         100.00
-------------------+------------------------------------------------

  Heterogeneity chi-squared =  20.16 (d.f. = 16) p = 0.213
  I-squared (variation in ES attributable to heterogeneity) =  20.6%
  Estimate of between-study variance Tau-squared =  0.0648

  Test of ES=1 : z=   5.29 p = 0.000
```

This difference in results is due to a) random error of the simulations and b) the different way the two commands handle studies without missing participants in one or both groups. More specifically, for these trials `metamiss2` assumes that the probability of observing the data is $\pi_{ij} = 1$, while `metamiss` assumes that the probability is not constant but a random variable.

An important advantage of `metamiss2` when using the bootstrap method is that it runs much faster (i.e. about 10 times) than `metamiss` due to coding in *mata*.

4.2. Pairwise meta-analysis, continuous data

The second example involves data from 8 trials that compare the effectiveness of mirtazapine versus placebo for major depression. The outcome is change in depression symptoms measured on a standardized rating scale (the HAMD21 scale).

```
. li,clean noo

    id         study      yp     sdp    np    mp      ym    sdm    nm    mm
     1   Claghorn1995    -11.4   10.2   19    26    -14.5   8.8    26    19
     2    MIR 003-003    -11.5    8.3   24    21      -14   7.3    27    18
     3    MIR 003-008    -11.4      8   17    13    -13.2     8    12    18
     4    MIR 003-020     -6.2    6.5   24    19      -13     9    23    21
     5    MIR 003-021    -17.4    5.3   21    29    -13.8   5.9    22    28
     6    MIR 003-024    -11.1    9.9   27    23    -15.7   6.7    30    20
     7    MIR 84023a     -11.9    8.6   33    24    -14.2   7.6    35    25
     8    MIR 84023b     -11.8    8.3   48    18    -14.7   8.4    51    13
```

We first describe the departure from MAR using the $IMDOM$. We assume a systematic departure from the MAR assumption where for the mirtazapine group $IMDOM$ has mean -0.5 with $sd(IMDOM) = 1$ and for the placebo group $IMDOM$ has mean 1 with $sd(IMDOM) = 1.5$. This means that we think it likely that missing participants had better outcomes than observed participants in the mirtazapine group (e.g. they left the study because of early response with important side effects) while the opposite is true in the placebo group (e.g. they left the study because of lack of efficacy). We also assume that $IMDOM$s are correlated between the two groups with $\rho = 0.5$ and we compare the results with ACA (i.e. when IMP=0 without uncertainty):

```
. metamiss2 nm mm ym sdm np mp yp sdp, impmean(-0.5 1) impsd(1 1.5) impcorr(0.5)
compare(impmean(0) impsd(0)) md metanopt(lcols(study))

Primary analysis


********************************************************************
******** METAMISS2: meta-analysis allowing for missing data *******
******** Informative missingness parameter with uncertainty *******
********************************************************************

Informative missingness parameter: IMDOM
Measure of interest:                Mean difference
Assumed distribution for IMP:       Experimental group ~ N(-0.5,1^2)
                                    Control group ~ N(1,1.5^2)
IMP correlation between groups:     0.5
Method for first stage model:       Taylor series approximation
Second stage model:                 Random effects meta-analysis

(Calling metan with options: lcols(study)...)


Secondary analysis
```

```
********************************************************************
******** METAMISS2: meta-analysis allowing for missing data ******
********          Available cases analysis           *******
********************************************************************

Informative missingness parameter: IMDOM
Measure of interest:              Mean difference
Method for first stage model:     Taylor series approximation
Second stage model:               Random effects meta-analysis

(Calling metan with options: lcols(study)...)

            Study    |    ES    [95% Conf. Interval]
--------------------+----------------------------------------------
    Primary analysis
1                    |   -3.889   -9.783    2.005
2                    |   -3.167   -7.653    1.319
3                    |   -2.533   -8.583    3.516
4                    |   -7.480  -12.143   -2.818
5                    |    2.740   -0.940    6.420
6                    |   -5.260   -9.860   -0.660
7                    |   -2.929   -6.956    1.097
8                    |   -3.274   -6.645    0.096
 Sub-total          |
  D+L pooled ES      |   -3.046   -5.264   -0.828
--------------------+----------------------------------------------
    Secondary analysis
9                    |   -3.100   -8.799    2.599
10                   |   -2.500   -6.814    1.814
11                   |   -1.800   -7.712    4.112
12                   |   -6.800  -11.305   -2.295
13                   |    3.600    0.251    6.949
14                   |   -4.600   -9.038   -0.162
15                   |   -2.300   -6.166    1.566
16                   |   -2.900   -6.191    0.391
 Sub-total          |
  D+L pooled ES      |   -2.382   -4.729   -0.035
--------------------+----------------------------------------------

Test(s) of heterogeneity:
            Heterogeneity  degrees of
                statistic    freedom    P    I-squared**   Tau-squared

Primary analysis    13.92         7      0.053    49.7%        4.9682
Secondary analysis  16.92         7      0.018    58.6%        6.5355
** I-squared: the variation in ES attributable to heterogeneity)


Significance test(s) of ES=0

Primary analysis     z=  2.69    p = 0.007
Secondary analysis   z=  1.99    p = 0.047
```

Next we change the IMP to the $IMROM$ and run a sensitivity analysis with $IMROM = 1$ on a range of different values for $sd(logIMROM)$ using the bootstrap method:

```
. metamiss2 nm mm ym sdm np mp yp sdp, md sensitivity b imptype(logimrom)

********************************************************************
******** METAMISS2: meta-analysis allowing for missing data ******
******** Informative missingness parameter with uncertainty ******
****    Sensitivity analysis assuming departures from MAR    *****
********************************************************************
```
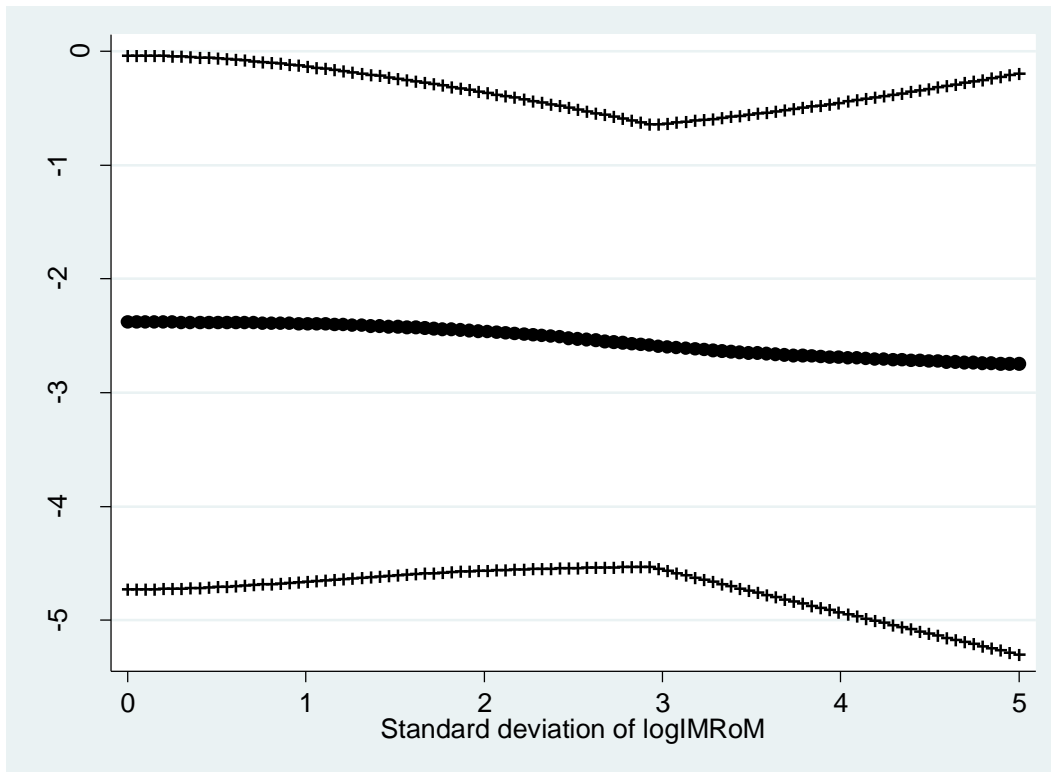
The graph shows that increasing the uncertainty of the IMP results into narrower confidence interval for the summery effect up to some point (~SD=3); this is related to the reduction of heterogeneity due to the extra variance introduced in the study-specific estimates. However, when very large uncertainty is assumed for IMP (SD>3), then this uncertainty is also reflected in the summary effect and therefore the confidence interval becomes wider.

4.3. Network meta-analysis

To exemplify the use of `metamiss2` in NMA we use a dataset that comprises a network of 12 trials comparing the effectiveness of 9 antidepressants. The outcome is again measured as the change score on the HAMD21 depression scale.

```
. li,clean noo

     id    t        y      sd     n      m
      1    8      4.6    8.65    39     13
      1    1        9    8.65    41      8
      2    1     7.56   12.31    39      1
      2    4    11.27   11.33    45      3
      3    6     14.2    11.1    75     18
      3    2     17.2    11.1    74     23
      4    2     13.5     2.1    12     15
      4    4     13.8     1.8    15     11
      5    4     7.76    2.89    21      3
      5    3     6.55    5.23    20      5
      6    4        7       5    45      0
      6    9        6       5    43      1
      7    9      8.4     5.4    67     11
      7    4      9.6     6.2    63     19
      8    4       11    9.12    55    145
      8    5     10.4    9.12    37    163
      9    5     15.8     3.3    30      3
      9    6     18.7     5.1    32      2
     10    6      8.2    7.52    55     20
     10    9      8.7    7.52    55     15
     11    6     10.4    7.52    15      3
```

```
11   7      7.4     7.52   16      2
12   9     11.3     8.65   91     37
12   6      12      8.65   89     32
```

Because of the complicated structure of data, `metamiss2` does not take arguments for the outcome when applied to NMA. Instead, the command `metamiss2` shall be executed after the data has been set up with the `network setup` command. This applies to any type of outcome that is handled with the network setup.

We first prepare the data in the 'augmented' format using the `network` package:

```
. network setup y sd n, trt(t) stud(id) nmiss(m) nocodes
{output omitted}
```

and then run `metamiss2` without arguments to obtain the ACA analysis:

```
. metamiss2
{output omitted}
```

To explore the impact of alternative assumptions we incorporate IMPs in our analysis. As in pairwise meta-analysis, IMPs can be treatment-specific. There are 9 treatments in the network and assumptions for the outcome among missing participants can be different depending on the administered treatment. Here, we consider that treatments 1, 2, 6 and 8 are associated with $IMDOM = 1$; for treatments 3, 4 and 9 $IMDOM = -1$ and $IMDOM = 0$ for treatments 5 and 7. We assume $sd(IMDOM) = 1$ for all treatments in the network. Additionally, drug-specific $IMDOMs$ can be correlated depending on the nature of missing data. Information about the pairwise correlation between the 9 $IMDOMs$ has to be collected in a matrix. In the matrix shown below, the correlation between the $IMDOMs$ for treatments 4 and 6 is $\rho_{4,6} = 0.5$ and between treatments 5 and 6 is $\rho_{5,6} = 0.2$:

$$C = \begin{pmatrix} 1 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 1 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 & 0.2 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.2 & 1 & 0.2 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 & 0.2 & 1 & 0.2 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.2 & 1 & 0.2 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.2 & 1 & 0.2 \\ 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.2 & 1 \end{pmatrix}$$

Note that here the choice of the correlation matrix is arbitrary but in practice it should be defined on the basis of expert opinion.

The matrix can be specified using the `mat` command

```
. mat C=J(9,9,0.5)+0.5*I(9)
forvalues i=4/8{
  mat C[`i',`=`i'+1']=0.2
  mat C[`=`i'+1',`i']=0.2
 }

. mat li C
```

```
symmetric C[9,9]
    c1  c2  c3  c4  c5  c6  c7  c8  c9
r1   1
r2  .5   1
r3  .5  .5   1
r4  .5  .5  .5   1
r5  .5  .5  .5  .2   1
r6  .5  .5  .5  .5  .2   1
r7  .5  .5  .5  .5  .5  .2   1
r8  .5  .5  .5  .5  .5  .5  .2   1
r9  .5  .5  .5  .5  .5  .5  .5  .2   1
```

To run the analysis, the `metamiss2` command has to take three arguments; the vector of the $IMDOMs$, the vector of their variances and matrix of correlations. We run the analysis using the bootstrap method:

```
. metamiss2, impmean(1 1 -1 -1 0 1 0 1 -1) impsd(1) impcorr(C) b

**********************************************************************
**** METAMISS2: network meta-analysis allowing for missing data ***
******** Informative missingness parameter with uncertainty *******
**********************************************************************

Informative missingness parameter: IMDOM
Measure of interest:               Mean difference
Assumed distribution for IMP:      1 ~ N(1,1^2)   (Reference group)
                                   2 ~ N(1,1^2)
                                   3 ~ N(-1,1^2)
                                   4 ~ N(-1,1^2)
                                   5 ~ N(0,1^2)
                                   6 ~ N(1,1^2)
                                   7 ~ N(0,1^2)
                                   8 ~ N(1,1^2)
                                   9 ~ N(-1,1^2)
IMP correlation between groups:    Matrix C
Method for first stage model:      Parametric Bootstrap (10000 draws)
Second stage model:                Random effects network meta-analysis

(Calling network meta with options: nograph ...)

Command is: mvmeta _y _S  , bscovariance(exch 0.5) longparm suppress(uv mm)  vars(_y_2
_y_3 _y_4 _y_5 _y_6 _y_7 _y_8 _y_9)
Note: using method reml
Note: using variables _y_2 _y_3 _y_4 _y_5 _y_6 _y_7 _y_8 _y_9
Note: 12 observations on 8 variables
Note: variance-covariance matrix is proportional to .5*I(8)+.5*J(8,8,1)

initial:       log likelihood = -50.831161
rescale:       log likelihood = -50.831161
rescale eq:    log likelihood = -50.427058
Iteration 0:   log likelihood = -50.427058
Iteration 1:   log likelihood =  -50.39345
Iteration 2:   log likelihood = -50.393443
Iteration 3:   log likelihood = -50.393443

Multivariate meta-analysis
Variance-covariance matrix = proportional .5*I(8)+.5*J(8,8,1)
Method = reml                             Number of dimensions    =     8
Restricted log likelihood = -50.393443    Number of observations  =    12
-----------------------------------------------------------------------------
             |     Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+---------------------------------------------------------------
_y_2         |
       _cons |   4.201509   2.568303     1.64   0.102    -.8322724    9.235291
-------------+---------------------------------------------------------------
_y_3         |
```

```
          _cons |    1.806593    2.776342     0.65    0.515    -3.634939     7.248124
--------------+----------------------------------------------------------------
_y_4          |
          _cons |    3.086682    2.430321     1.27    0.204    -1.676659     7.850024
--------------+----------------------------------------------------------------
_y_5          |
          _cons |    .5450634    2.709961     0.20    0.841    -4.766362     5.856489
--------------+----------------------------------------------------------------
_y_6          |
          _cons |    2.808598    2.601193     1.08    0.280    -2.289646     7.906842
--------------+----------------------------------------------------------------
_y_7          |
          _cons |   -.3568005    3.786236    -0.09    0.925    -7.777686     7.064085
--------------+----------------------------------------------------------------
_y_8          |
          _cons |   -4.303123    1.982229    -2.17    0.030     -8.18822    -.4180255
--------------+----------------------------------------------------------------
_y_9          |
          _cons |    2.046148    2.517793     0.81    0.416    -2.888636     6.980932
--------------+----------------------------------------------------------------

Estimated between-studies SDs and correlation matrix:
             SD      _y_2      _y_3      _y_4      _y_5      _y_6      _y_7      _y_8          _y_9
_y_2   3.837e-11       1         .         .         .         .         .         .             .
_y_3   3.837e-11      .5         1         .         .         .         .         .             .
_y_4   3.837e-11      .5        .5         1         .         .         .         .             .
_y_5   3.837e-11      .5        .5        .5         1         .         .         .             .
_y_6   3.837e-11      .5        .5        .5        .5         1         .         .             .
_y_7   3.837e-11      .5        .5        .5        .5        .5         1         .             .
_y_8   3.837e-11      .5        .5        .5        .5        .5        .5         1             .
_y_9   3.837e-11      .5        .5        .5        .5        .5        .5        .5             1
mvmeta command stored as F9
```

Accounting for missing outcome data in this particular example had little impact on the results, which might be due to the arbitrary assumptions we made about the IMPs. Treatment 8 appears to be more effective than treatment 1, as in the ACA analysis. The confidence intervals of all relative effects are slightly narrower compared to ACA, while heterogeneity was estimated to be near zero.

5. Discussion

metamiss2 and metamiss are almost equivalent for meta-analyses with binary outcome data, and they give identical answers when the Taylor series method is used to account for uncertainty. However, small discrepancies exist between the two commands. First, metamiss has the option to perform analyses of missing dichotomous data based on *reasons for missingness* (White and Higgins, 2009). This approach allows different assumptions to be made within each study group at the patient level and not only on average as metamiss2 (Higgins et al., 2008). However, patient-level information about reasons for missingness is the exception rather than the norm in a meta-analysis of aggregated data. Second, the option to use the *Gauss-Hermite quadrature* estimation method is not available in metamiss2. However, the parametric bootstrap method in metamiss2 is very fast and thus it can be used routinely as an alternative to quadrature. Note that the Monte Carlo method which is available in metamiss is fully Bayesian and thus can show small numerical differences from the parametric bootstrap method in metamiss2.

A limitation of metamiss2 is that finite-sample correction for SMD has not been incorporated in the present code; this correction allows for uncertainty in the observed study-specific standard deviations when trial sample sizes are small. Future work will explore the potential

to enable an assumption that IMPs are correlated across different studies (White et al., 2008b, p. 2).

There is no unique best approach to handle missing outcome data in meta-analysis. ACA is usually a sensible starting point and will often be the primary analysis. Since the IMP parameters cannot be estimated from the observed data, values most be given to them based on judgement and on evidence external to the meta-analysis. Thus, sensitivity analyses using different plausible values of IMPs are necessary to assess the robustness of results to different assumptions about the missing data. The sensitivity option in `metamiss2` sets the IMP means and correlation to zero and gradually increases the IMP standard deviations. This reflects a minor departure from MAR. In practice, we would expect the IMP mean to be non-zero. We may conduct additional sensitivity analyses changing the value of both mean and SD (one at a time) of IMP parameters and assuming each of them common and different across groups, and monitor how sensitive results are to these changes. Other sensitivity analysis strategies were suggested by White et al (2008a). In all cases, discussion with subject matter experts is needed to choose sensible distributions for the IMPs.

6. References

Chaimani, A., Salanti, G., 2015. Visualizing assumptions and results in network meta-analysis: The network graphs package. Stata J. 15, 905–950.

Harris, R., Bradburn, M., Deeks, J., Harbord, R., Altman, D., Sterne, J., 2008. Metan: fixed- and random-effects meta-analysis. Stata J. 8, 3.

Higgins, J.P., White, I.R., Wood, A.M., 2008. Imputation methods for missing outcome data in meta-analysis of. Clin. Trials 5, 225–239. doi:10.1177/1740774508091600

Little, R.J., Rubin, D.B., 2014. Statistical analysis with missing data. John Wiley & Sons.

Mavridis, D., White, I.R., Higgins, J.P.T., Cipriani, A., Salanti, G., 2015. Allowing for uncertainty due to missing continuous outcome data in pairwise and network meta-analysis. Stat. Med. 34, 721–741. doi:10.1002/sim.6365

Palmer, T.M., Sterne, J.A.C., 2016. Meta-analysis in Stata: An updated collection from the Stata Journal, Second. ed. Stata Press.

Salanti, G., Del Giovane, C., Chaimani, A., Caldwell, D.M., Higgins, J.P.T., 2014. Evaluating the Quality of Evidence from a Network Meta-Analysis. PLoS ONE 9. doi:10.1371/journal.pone.0099682

Salanti, G., Higgins, J.P., Ades, A.E., Ioannidis, J.P., 2008. Evaluation of networks of randomized trials. Stat Methods MedRes 17, 279–301. doi:10.1177/0962280207080643

Spineli, L.M., Higgins, J.P., Cipriani, A., Leucht, S., Salanti, G., 2013. Evaluating the impact of imputations for missing participant outcome data in a network meta-analysis. Clin.Trials. doi:10.1177/1740774512470317

White, I.R., 2015. Network meta-analysis. Stata J. 15, 951–985.

White, I.R., Higgins, J.P., 2009. Meta-analysis with missing data. Stata J. 9, 57–69.

White, I.R., Higgins, J.P.T., Wood, A.M., 2008a. Allowing for uncertainty due to missing data in meta-analysis--part 1: two-stage methods. Stat. Med. 27, 711–727. doi:10.1002/sim.3008

White, I.R., Thomas, J., 2005. Standardized mean differences in individually-randomized and cluster-randomized trials, with applications to meta-analysis. Clin. Trials 2, 141–151. doi:10.1191/1740774505cn081oa

White, I.R., Welton, N.J., Wood, A.M., Ades, A.E., Higgins, J.P.T., 2008b. Allowing for uncertainty due to missing data in meta-analysis—Part 2: Hierarchical models. Stat. Med. 27, 728–745. doi:10.1002/sim.3007

7. About the authors

Anna Chaimani is a Junior Chair at Paris Descartes University, France. Her research interests focus on methodology for pairwise and network meta-analysis. She is the author of `network graphs` for network meta-analysis.

Ian White is Professor of Statistical Methods for Medicine at the MRC Clinical Trials Unit at UCL, UK. His research interests focus on handling missing data, noncompliance and measurement error in the analysis of clinical trials, observational studies, and meta-analysis. He is the author of `mvmeta` for multivariate meta-analysis and `network` for network meta-analysis.

Dimitris Mavridis is an assistant professor in the Department of Primary Education in the University of Ioannina, Greece. His research interests focus on methodology for pairwise and network meta-analysis and more specifically on methods for accounting for missing outcome data and publication bias.

Julian Higgins is Professor of Evidence Synthesis at the University of Bristol, UK. His research interests are around statistical and non-statistical methods for systematic reviews, meta-analysis and research synthesis more generally.

Georgia Salanti is Associate Professor of Biostatistics and Epidemiology at the University of Bern, Switzerland. Her research interests focus on methods for evidence synthesis.