Allowing for uncertainty due to missing outcome data in meta-analysis

Anna Chaimani

Research Center of Epidemiology and Statistics Sorbonne Paris Cité (CRESS-UMR1153)
Paris Descartes University

Dimitris Mavridis

Department of Primary Education
University of Ioannina, Greece

With much help from : Georgia Salanti, Ian White, Julian Higgins

Scope of this webinar

- To present a “new” method to deal with missing outcome data in aggregate data meta-analysis

- It is not about
  - missing studies (publication bias)
  - missing outcomes (selective outcome reporting)
  - missing statistics (standard deviations/errors)

- It is not about IPD meta-analysis

Outline

- Why are missing data a problem in meta-analysis?

- Examples
  - haloperidol in schizophrenia (dichotomous outcome)
  - mirtazapine in depression (continuous outcome)

- Very brief overview of currently used methods

- Proposed method based on Informative Missingness Parameters (IMP)

Why missing outcome data matter?

- Missing outcome data are common in RCT's
  - In mental health, the dropout rate may exceed 50%

- It creates two main problem at RCT level:
  - loss in power and precision
    - because the sample size decreases
  - bias (maybe)
    - any analysis must make an untestable assumption about missing data
    - wrong assumption → biased estimates

- There is no remedy for missing data - we can only do sensitivity analyses and see how much the results change under different assumptions

Fictional example

Study	Observed	SMD	Naïve SE (relative weight)	Randomized
1	100	0	0.07 (20%)	100
2	100	0.1	0.07 (20%)	120
3	100	0.2	0.07 (20%)	150
4	100	0.3	0.07 (20%)	200
5	100	0.4	0.07 (20%)	300

- Would you give each study the same weight in the meta-analysis?
  a) Yes, because for all studies we have data for 100 participants
  b) No, I will give the largest weight to study 5 because has the largest number of randomized participants
  c) No, I will give the largest weight to study 1 because observed and randomized participants are the same
  d) No, I will give the smallest weight to study 1 because the SMD is 0

Meta-analysis with studies with missing outcome data

- There are methods to address missing outcome data at trial level (LOCF, multimple imputation, etc.)
  - –NOT addressed here as you need IPD
- What do you do when you have aggregate data from studies with missing data?
  - – any meta-analysis makes an untestable assumption about missing data even if reviewers don't realize it!

Missing data mechanisms

- Missing completely at random (MCAR)

The probability that data are missing does not depend neither on observed nor on unobserved data

– In an RCT of antihypertensives that measures blood pressure (BP), some data are missing due to the breakdown of an automatic sphygmomanometer

- Missing at random (MAR)

The probability that data are missing does not depend on the outcome or unobserved factors that impact on the outcome

– In an RCT of antihypertensives that measures blood pressure (BP), older participants are more likely to have their BP recorded. Missing data are MAR if at any age, individuals with low and high BP are equally likely to have their BP recorded

- Missing not at random (MNAR)

The probability that data are missing depends on the outcome

– In an RCT of antipsychotics individuals with relapse are more likely to leave the study early in the placebo group

Meta-analysis of studies with missing outcome data

Outcome is reponse to treatment in some standardized scale

- Example: Haloperidol vs. placebo in schizophrenia
  - 17 trials in a Cochrane Review
  - We use an inverse-variance-weighted random-effect analysis on the risk ratio scale

- Example: Mirtazapine vs. placebo in depression
  - 8 trials taken from a recent network meta-analysis (Cipriani et al, 2018)
  - We use an inverse-variance-weighted random-effect analysis on the standardized mean difference scale

Haloperidol vs. placebo in schizophrenia (Beasley 1998)

	Success	Failure	Missing
Haloperidol	29	18	22
Placebo	20	14	34

- Outcome: clinical global improvement (yes/no)
- RR=1.05 (95% CI: 0.73-1.50)
- Missing rates are 32% for haloperidol and 50% for placebo

Haloperidol vs. placebo in schizophrenia (Beasley 1998)

	Success	Failure	Missing
Haloperidol	29	18	22
Placebo	20	14	34

- What is the data missing mechanism?
  a) MAR, because data are missing from both groups
  b) Most probably MAR, because data are missing from both groups
  c) Most probably MNAR, because this is the psychiatrists' opinion
  d) We are clueless

Haloperidol vs. placebo in schizophrenia (Beasley 1998)

	Success	Failure	Missing
Haloperidol	29	18	22
Placebo	20	14	34

- How would you analyze these data?
  a) I would analyze only the completers
  b) I would assume missing participants did not respond to treatment
  c) I would assume the same risks in the missing participants as those in the observed
  d) I would exclude the study from the meta-analysis
  e) None of the above

Haloperidol vs. placebo in schizophrenia (Beasley 1998)

	Success	Failure	Missing
Haloperidol	29	18	22
Placebo	20	14	34

Success rates: 29/47=0.62 vs 20/34=0.59 (Available Cases Analysis, ACA)

RR=1.05 (95% CI: 0.73-1.50)

Which i

ANY analysis makes assumptions which, if wrong, produces biased results!

Succes

RR=1.43 (95% CI: 0.90-2.27)

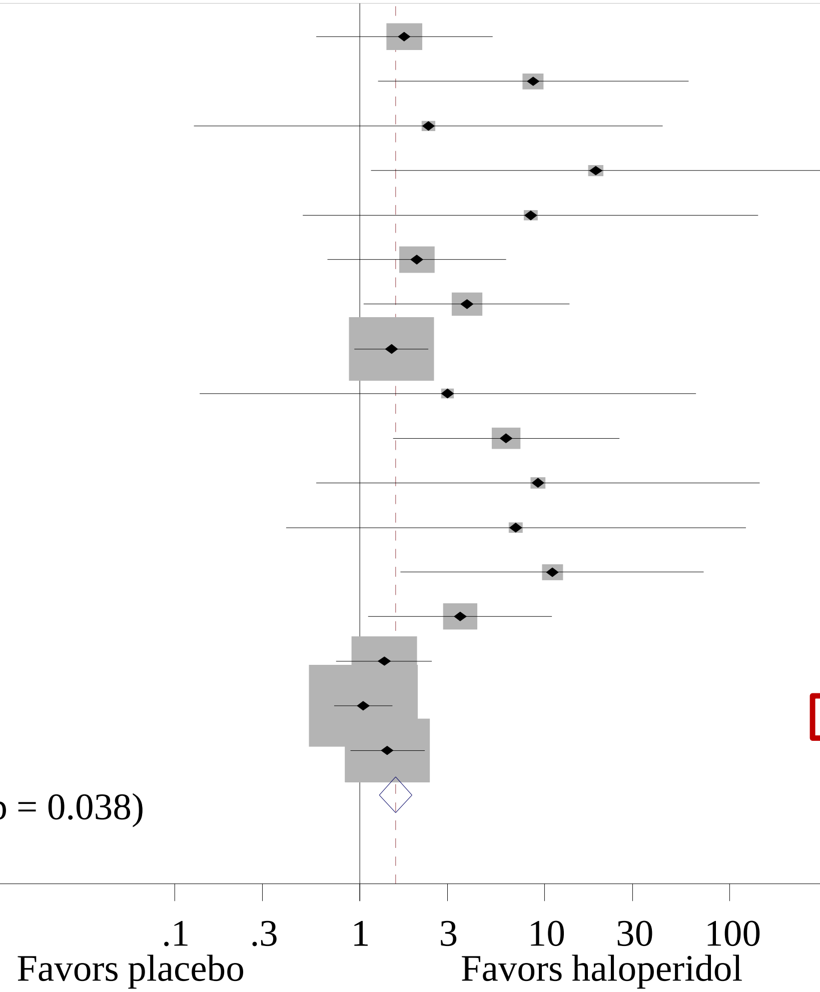
Which is the assumption behind? ***Successes have no chance to dropout!***

Haloperidol vs. placebo in schizophrenia

	Haloperidol			Placebo		
	rh	fh	mh	rp	fp	mp
Arvanitis	25	25	2	18	33	0
Beasley	29	18	22	20	14	34
Bechelli	12	17	1	2	28	1
Borison	3	9	0	0	12	0
Chouinard	10	11	0	3	19	0
Durost	11	8	0	1	14	0
Garry	7	18	1	4	21	1
Howard	8	9	0	3	10	0
Marder	19	45	2	14	50	2
Nishikawa 82	1	9	0	0	10	0
Nishikawa 84	11	23	3	0	13	0
Reschke	20	9	0	2	9	0
Selman	17	1	11	7	4	18
Serafetinides	4	10	0	0	13	1
Simpson	2	14	0	0	7	1
Spencer	11	1	0	1	11	0
Vichaiya	9	20	1	0	29	1

Haloperidol vs. placebo
Available case analysis

author	RR (95% CI)	% Weight
Garry	1.75 (0.58, 5.24)	3.37
Durost	8.68 (1.26, 59.95)	1.09
Simpson	2.35 (0.13, 43.53)	0.48
Vichaiya	19.00 (1.16, 311.96)	0.52
Serafetinides	8.40 (0.50, 142.27)	0.51
Howard	2.04 (0.67, 6.21)	3.27
Reschke	3.79 (1.06, 13.60)	2.48
Selman	1.48 (0.94, 2.35)	19.11
Nishikawa_82	3.00 (0.14, 65.90)	0.42
Bechelli	6.21 (1.52, 25.35)	2.05
Nishikawa_84	9.20 (0.58, 145.76)	0.53
Borison	7.00 (0.40, 122.44)	0.49
Spencer	11.00 (1.67, 72.40)	1.14
Chouinard	3.49 (1.11, 10.95)	3.10
Marder	1.36 (0.75, 2.47)	11.37
$\boxed{\text{Beasley}}$	1.05 (0.73, 1.50)	31.22
Arvanitis	1.42 (0.89, 2.25)	18.86
Overall (I-squared = 41.4%, p = 0.038)	1.57 (1.28, 1.92)	100.00



missing rate 41%

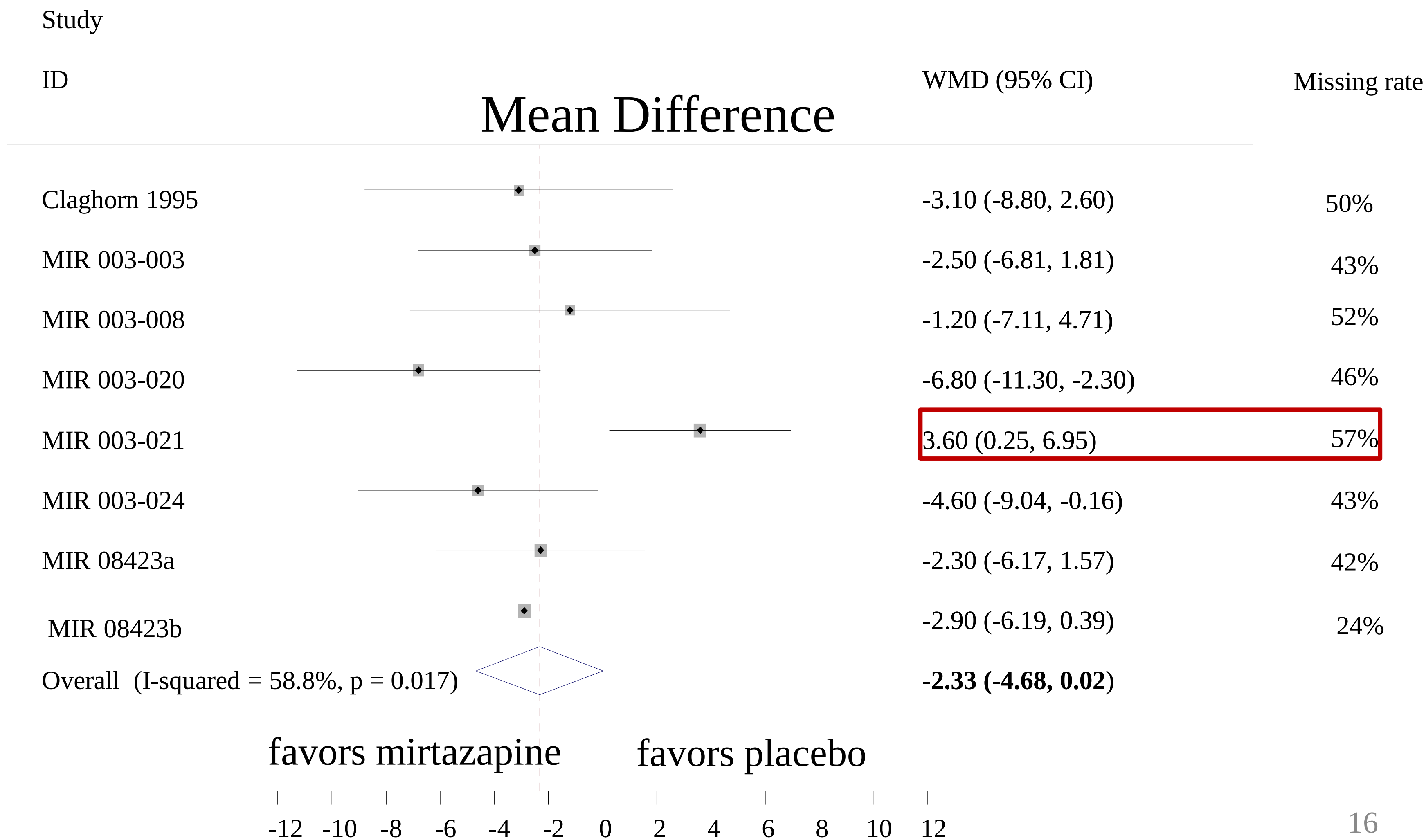
Mirtazapine vs Placebo for depression

Change in depression symptoms measured on the HAMD21 scale

Study	Placebo				Mirtazapine			
	xp	**sdp**	**n**	**m**	**xm**	**sdm**	**n**	**m**
Claghorn 1995	-11.4	10.2	19	26	-14.5	8.8	26	19
MIR 003-003	-11.5	8.3	24	21	-14	7.3	27	18
MIR 003-008	-11.4	8	17	13	-13.2	8	12	18
MIR 003-020	-6.2	6.5	24	19	-13	9	23	21
MIR 003-021	-17.4	5.3	21	29	-13.8	5.9	22	28
MIR 003-024	-11.1	9.9	27	23	-15.7	6.7	30	20
MIR 84023a	-11.9	8.6	33	24	-14.2	7.6	35	25
MIR 84023b	-11.8	8.3	48	18	-14.7	8.4	51	13

Mirtazapine vs Placebo for depression

Complete case analysis



Summary of possible analyses (Cochrane Handbook)

Analysis	Outcome	Description of method/how it handles missing participants	Assumptions about missing outcome data	Adequacy for addressing missing data
Available cases	binary continuous	ignore them	a random sample of all participants	valid under missing at random (MAR)
worst (best)-case scenario	binary	imputes failures in the treatment group and successes in the control (or vice-versa)	worse in the experimental group (better in the experimental group)	*inflates sample size and erroneously increase precision/reduce standard errors*
mean imputation	continuous	imputes the mean value	the same as observed	
other simple imputation	binary continuous	imputes specific number of successes/mean value	explicit assumptions about them	
gamble-hollis	binary	downweight studies according to best/worst case scenarios	studies with large differences between best/worst case scenario are less reliable	too extreme downweighting
The suggested model	binary continuous	downweight studies with high missing rates	the more the missing rate the less reliable is the estimate	**Accounts for uncertainty in the missing outcome data - Expert opinion can also be used**

A general approach

- We propose the informative missingness parameter (IMP) as a general way to think about missing data
- **Definition**: IMP relates a summary statistic in the missing group to the corresponding summary statistic in the observed group
- **IMOR** (Informative missing odds ratio): the odds of success in the missing group over the odds of success in the observed group
- **IMDOM** (Informative missing difference of means): the mean in the missing group minus the mean in the observed group
- IMPs may different between intervention and control arm
- IMPs are not known, but we can suggest clinically plausible values

Characteristics of our approach

- We don't impute missing data!
- We simply make assumptions about the outcome in the missing data and how it is related to the outcome in the observed data (we "update" the study outcome using the observed data)
- We employ a pattern-mixture model
- In the entire procedure we account for the fact that data is not fully observed
  - **This is very important in order to obtain correct standard errors from studies! (see later…)**

Pattern mixture model

i refers to study, *j* refers to arm, *k* refers to individual

$$Y = \left(Y^{obs}, Y^{miss}\right)'$$

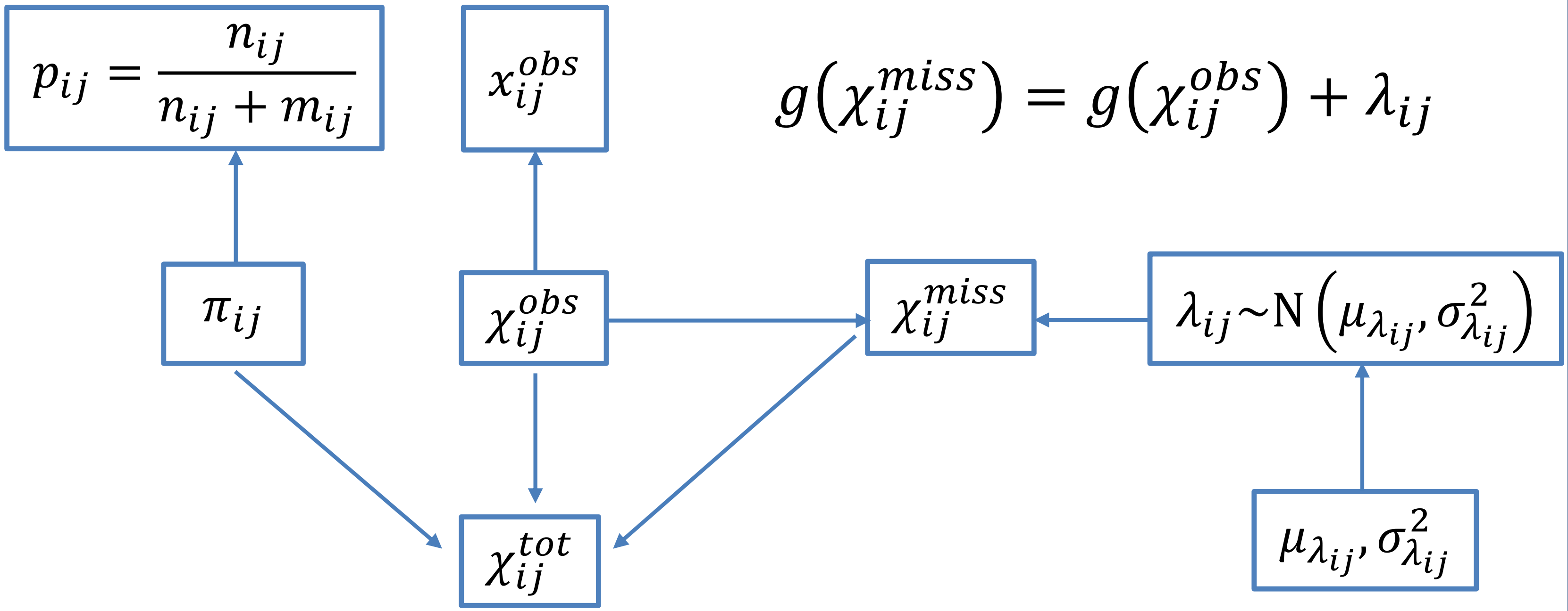
$$R_{ijk} = \begin{cases} 1\ if\ outcome\ is\ reported \\ 0\ if\ outcome\ is\ missing \end{cases} \qquad P\left(R_{ijk} = 1\right) = \pi_{ij}^{obs}$$

$$E\left(Y_{ijk} | R_{ijk} = 1\right) = \chi_{ij}^{obs} \qquad E\left(Y_{ijk} | R_{ijk} = 1\right) = \chi_{ij}^{miss}$$

$$f(Y, R) = f(Y|R) f(R)$$

Pattern mixture model

Model for arm *j* of study *i*



$$\chi_{ij}^{tot} = p_{ij}\chi_{ij}^{obs} + (1 - p_{ij})\chi_{ij}^{miss}$$

$$\chi_{ij}^{tot} = p_{ij}\chi_{ij}^{obs} + \left(1 - p_{ij}\right)g^{-1}\left(g\left(\chi_{ij}^{obs}\right) + \lambda_{ij}\right)$$

Continuous outcome
Informative missingness difference of means

$$g\left(\chi_{ij}^{miss}\right) = \lambda_{ij} + g\left(\chi_{ij}^{obs}\right)$$

If g is the identity function

$$\lambda_{ij} = \chi_{ij}^{miss} - \chi_{ij}^{obs}$$

IMP =λ= mean in missing – mean in observed

- **λ=1** the mean in the missing participants exceed the mean in the observed participants by one unit
- **λ=-1** the mean in the missing participant is one unit less compared to the mean of the observed participants
- λ=0 the data is missing at random

Dichotomous outcome

Informative missingness odds ratio

$$g\left(\chi_{ij}^{miss}\right) = \lambda_{ij} + g\left(\chi_{ij}^{obs}\right)$$

If g is the identity function

$$\lambda_{ij} = \chi_{ij}^{miss} - \chi_{ij}^{obs}$$

IMP =λ= log(odds) in missing – log(odds) in observed

$$\text{IMOR}=\exp(\lambda)= \frac{\text{odds of success in missing}}{\text{odds of success in observed}}$$

Dichotomous outcome
Informative missingness odds ratio

$$\text{IMOR} = \frac{\text{odds of success in missing}}{\text{odds of success in observed}}$$

- **IMOR=2** the odds of success is double in the missing participants rather than the observed

  (e.g. maybe people leave the study because of early response!)

- **IMOR=0.5** the odds of success is half in the missing participants rather than the observed

  (e.g. maybe people leave the study because of they are disappointed as they don't see any improvement!)

- **IMOR=1 the data is missing at random**

  Note that in this case $\lambda = log(IMOR)$, hence $\lambda = 0$ implies MAR

Assumptions about the IMP

- Missing at random $\lambda_{ij} = 0$

- Free $\lambda_{ij} \sim N\left(\mu_{\lambda_{ij}}, \sigma^2_{\lambda_{ij}}\right)$

- Study-specific $\lambda_i \sim N\left(\mu_{\lambda_i}, \sigma^2_{\lambda_i}\right)$

- Correlated λ's

$$\begin{pmatrix} \lambda_{iC} \\ \lambda_{iT} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_{\lambda_{iC}} \\ \mu_{\lambda_{iT}} \end{pmatrix}, \begin{pmatrix} \sigma^2_{\lambda_{iC}} & \rho_\lambda \sigma_{\lambda_{iC}} \sigma_{\lambda_{iT}} \\ \rho_\lambda \sigma_{\lambda_{iC}} \sigma_{\lambda_{iT}} & \sigma^2_{\lambda_{iT}} \end{pmatrix}\right)$$

A two-stage approach

- At the first-stage we compute an adjusted effect size

$$\beta_i = f\left(x_{iT}^{tot}\right) - f\left(x_{iC}^{tot}\right)$$

  – If the outcome is dichotomous (x is the risk of the event)

    - β_i is the Risk Difference if f is the identity function
    - β_i is the logarithm of Risk Ratio if f is the logarithmic function
    - β_i is the logarithm of Odds Ratio if f is the logit function

  – If the outcome is continuous (x is the mean outcome)

    - β_i is the Mean Difference if f is the identity function
    - β_i is the Standardized Mean Difference if f is the identity function divided with the pooled standard deviation

- At the second-stage we compute an inverse-variance random effects meta-analysis

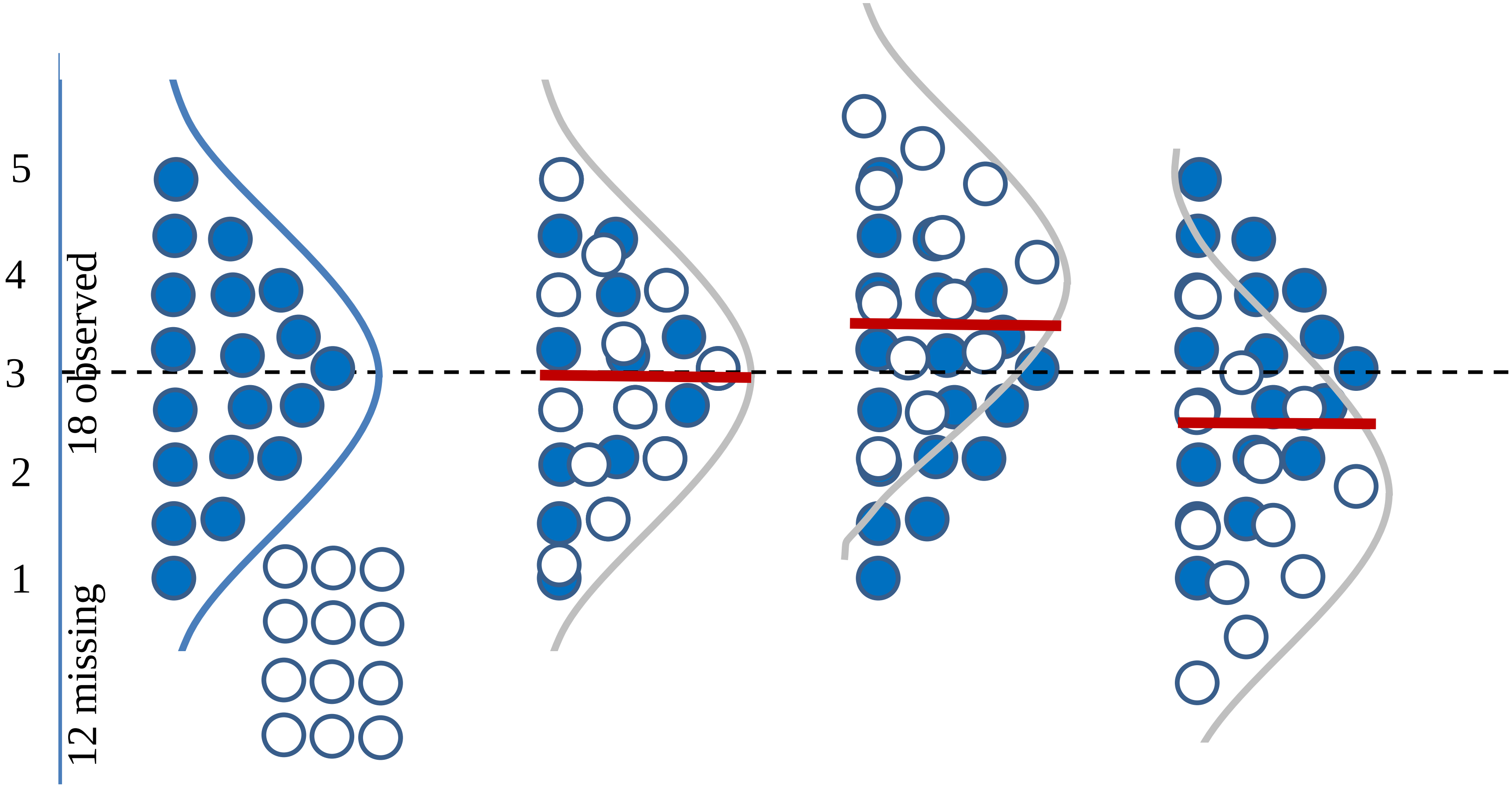
Haloperidol vs placebo in schizophrenia

Study	Haloperidol			Placebo		
	rh	fh	mh	rp	fp	mp
Beasley	29	18	22	20	14	34
Selman	17	1	11	7	4	18
Marder	19	45	2	14	50	2

Let us assume:
IMOR=1 in Haloperidol (MAR)
IMOR=0.5 in Placebo (IM due to lack of effectiveness)

Study	Haloperidol (IMOR 1)			Placebo (IMOR 0.5)			OR
	Odds observed	Odds missing	Total odds	Odds observed	Odds missing	Total odds	
Beasley	1.61	1.61	1.61	1.43	0.72	1.01	1.60
Selman	17	17	17	1.75	0.88	1.13	15.04
Marder	0.42	0.42	0.42	0.28	0.14	0.27	1.56

Meta-analyze these!
(you need their SEs… see later)



mean of observed	3	3	3
mean of missing	3	4	2.6
total mean	3	3.4	2.6
	(3×18+ 3×12)/30	(3×18+ 4×12)/30	(3×18+ 2×12)/30
IMDOM	0	1	-1

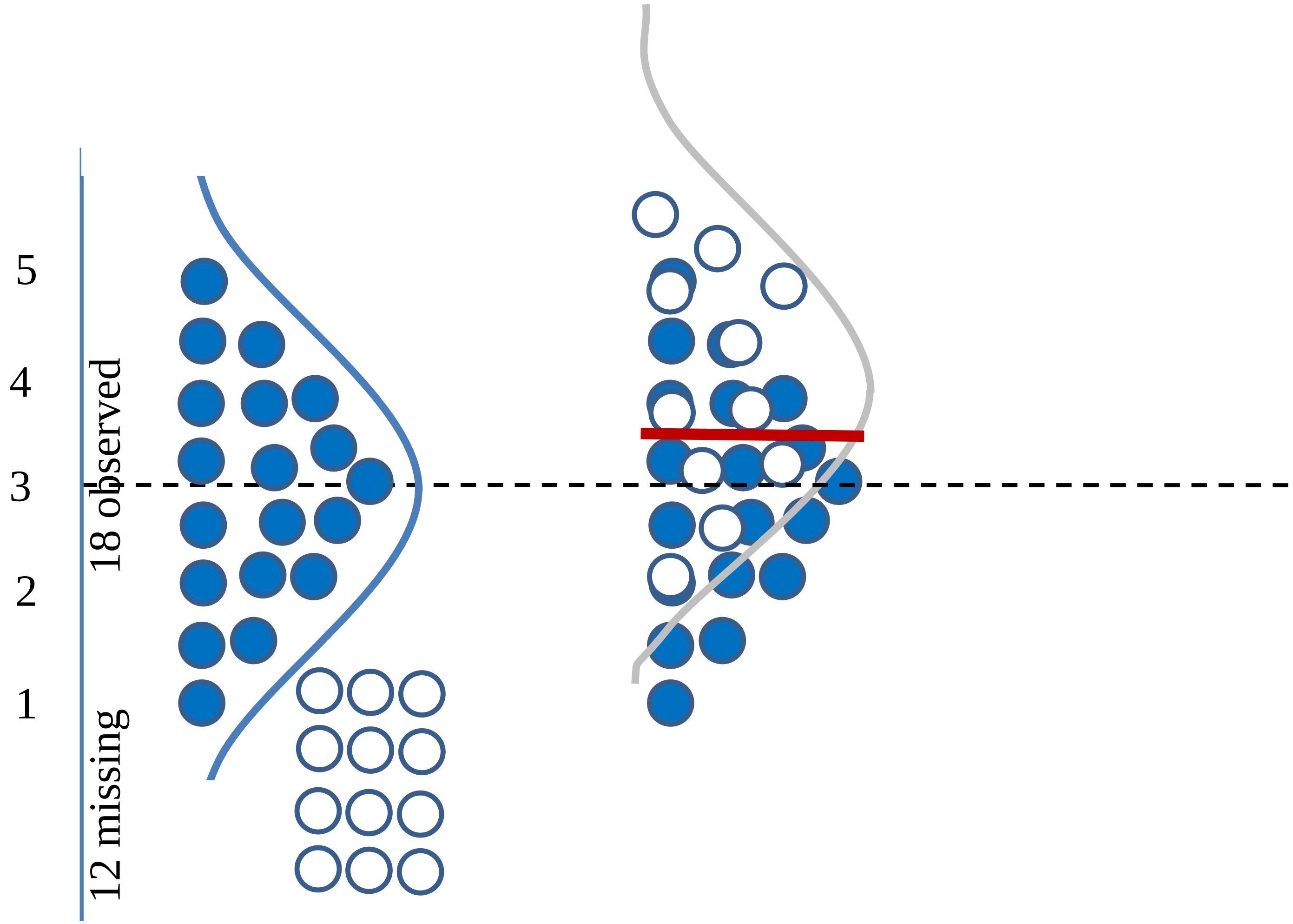
Work out the total means starting from IMDOM

- We ask a clinician (or several!) with experience in clinical trials in the field:

  "Out of 100 patients randomized in drug X, 60 finished the study and had a mean score 3 whereas 40 patients did not finish. What do you guess would be the mean score in those who did not finish?"

  -he answered "the mean score in those who did not finish is on average 4"

- What is the value of IMDOM?
  a) IMDOM=7
  b) IMDOM=4
  c) IMDOM=3
  d) IMDOM=1



Observed mean= 3	Observed mean = 3
Mean of missing	Observed mean +IMDOM =4
Total mean	3.4 (3×18+ 4×12)/30
IMDOM λ	1

Mirtazapine vs placebo for depression

Study	Placebo				Mirtazapine			
	xp	sdp	n	m	xm	sdm	n	m
Claghorn 1995	-11.4	10.2	19	26	-14.5	8.8	26	19
MIR 003-003	-11.5	8.3	24	21	-14	7.3	27	18
MIR 003-008	-11.4	8	17	13	-13.2	8	12	18

We assume **IMDOM=1 for Placebo** (the symptoms decreased in the missing participants) and **IMDOM=-1 for Mirtazapine** (missing participants left because of early response)

Study	Placebo		Mirtazapine		MD
	Missing mean	Total mean	Missing mean	Total mean	
Claghorn 1995	-10.4	-10.82	-15.5	-14.92	-4.10
MIR 003-003	-10.5	-11.03	-15	-14.40	-3.37
MIR 003-008	-10.4	-10.97	-14.2	-13.80	-2.83

Meta-analyze these!
(you need their SEs… see later)

Fictional example

Study	Observed	SMD	Naïve SE (relative weight)	Randomized
1	100	0	0.07 (20%)	100
2	100	0.1	0.07 (20%)	120
3	100	0.2	0.07 (20%)	150
4	100	0.3	0.07 (20%)	200
5	100	0.4	0.07 (20%)	300

Would you give each study the same weight?

No, because uncertainty should be larger when you have more missing data!
The assumption on how the outcome in the missing and observed participants is related (λ_{ij}) has more impact on study 5 rather than on study 2!

The observed sample size is not the only source of uncertainty!

First source of extra uncertainty: Proportion of missing data!

Fictional example

Studies with means and same standard deviations per arm and observed sample size, but different missingness.

Study	Observed	SMD	Naïve SE (relative weight)	Randomized
1	100	0	0.07 (20%)	100
2	100	0.1	0.07 (20%)	120
3	100	0.2	0.07 (20%)	150
4	100	0.3	0.07 (20%)	200
5	100	0.4	0.07 (20%)	300

Let us assume that there are no differences in the mean outcome between missing and observed participants

- We can NEVER be sure that the mean in the missing is exactly the same as in the observed
- We have some **uncertainty as to what exactly is the mean in the missing data**
- This can be represented by uncertainty in IMDOM!
- **Let us assume assume IMDOM=0 with uncertainty interval (-1, 1) – that is approximately translated to** $\boldsymbol{\lambda_{ij} \sim N(0, 0.5^2)}$

Second source of extra uncertainty: Uncertainty about the assumption for IMP

IMOR/IMDOM have uncertainty

IMDOM=0, (0.5, 0.0)

" the missing participants are likely to have on average the same score as the observed participant with uncertainty interval (0.5, 0.5)"

Equivalently we need to set values for IMDOM and sd(IMDOM)

$$\boldsymbol{\lambda_{ij} \sim N(0, 0.5^2)}$$

Similarly for IMOR…

IMOR=1, (0.6, 1.6)

" the missing participants are likely to have the same odds of success as the observed participant with uncertainty interval (0.6, 1.6)"

Equivalently we need to set values for IMOR and sd(logIMOR)

$$\boldsymbol{\lambda_{ij} \sim N(0, 0.25^2)}$$

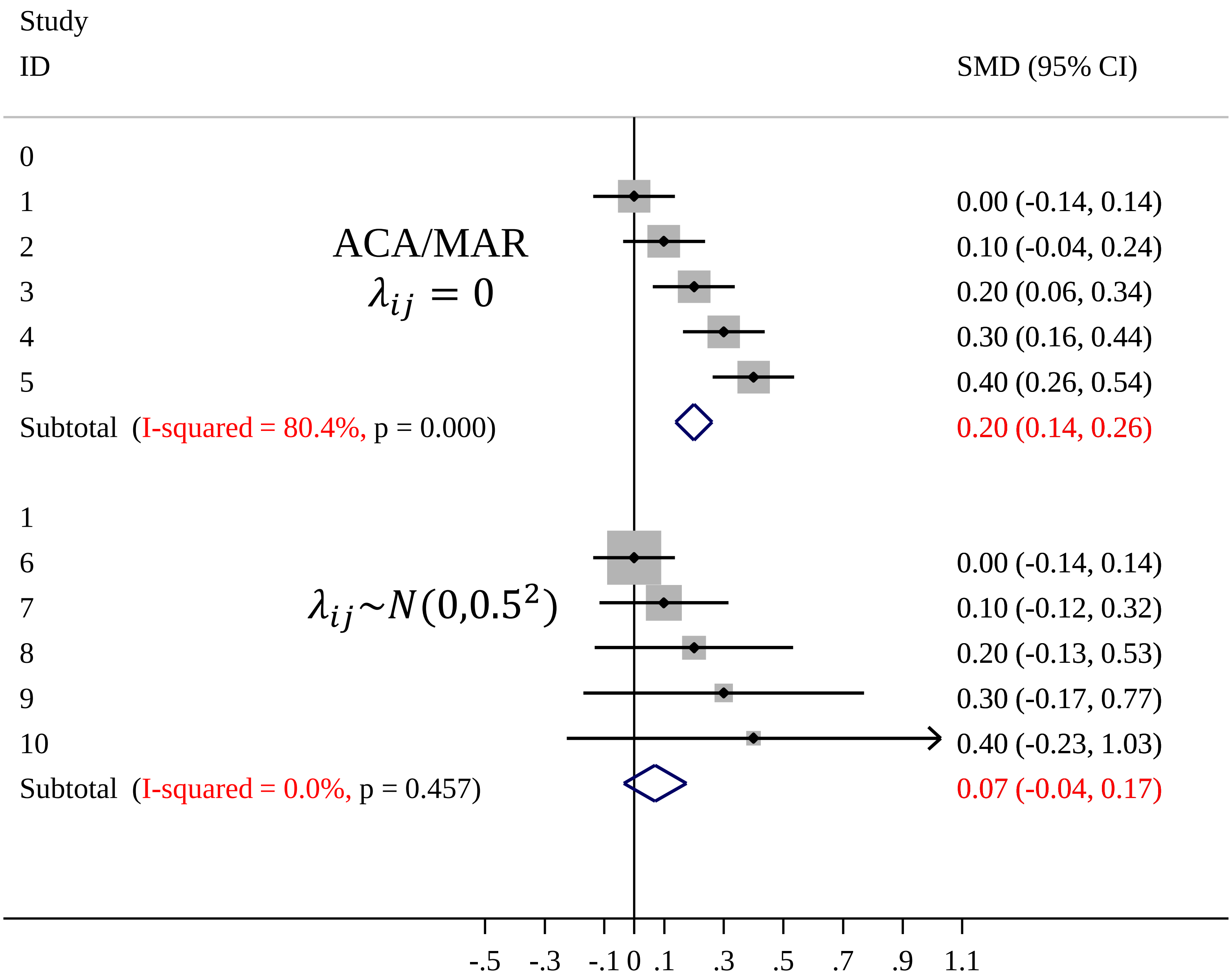
Fictional example

Studies with means and same standard deviations per arm and observed sample size, but different missingness.

Study	Observed	SMD	Naïve SE (relative weight)	Randomized	Corrected SE (relative weight)
1	100	0	0.07 (20%)	100	0.07 (59%)
2	100	0.1	0.07 (20%)	120	0.11 (24%)
3	100	0.2	0.07 (20%)	150	0.17 (10%)
4	100	0.3	0.07 (20%)	200	0.24 (5%)
5	100	0.4	0.07 (20%)	300	0.32 (3%)

We assume IMDOM=0 with uncertainty interval (-1, 1)

Studies with more missing data get less weight!


Study
ID
SMD (95% CI)
0
1
2
3
4
5
ACA/MAR
$\lambda_{ij} = 0$
0.00 (-0.14, 0.14)
0.10 (-0.04, 0.24)
0.20 (0.06, 0.34)
0.30 (0.16, 0.44)
0.40 (0.26, 0.54)
Subtotal (I-squared = 80.4%, p = 0.000)
0.20 (0.14, 0.26)
1
6
7
8
9
10
$\lambda_{ij} \sim N(0, 0.5^2)$
0.00 (-0.14, 0.14)
0.10 (-0.12, 0.32)
0.20 (-0.13, 0.53)
0.30 (-0.17, 0.77)
0.40 (-0.23, 1.03)
Subtotal (I-squared = 0.0%, p = 0.457)
0.07 (-0.04, 0.17)
-.5 -.3 -.1 0 .1 .3 .5 .7 .9 1.1

Key thing: Estimation of SE of the effect size

- To estimate SE(logRR), SE(logOR) and SE(SMD) you need mathematical manipulations or simulations (rather cumbersome!)
- Likely, Stata will do the trick for you!
  - Using Monte Carlo
  - Using a Taylor series approximation

For all mathematical details see:

- For continuous outcomes

Mavridis D., White I., Higgins J., Cipriani A., Salanti G **Allowing for uncertainty due to missing continuous missing outcome data in pairwise and network meta-analysis.** *Statistics in Medicine* 2014, *34*, *721–741*

- For dichotomous outcomes

White IR, Higgins JPT, Wood AM: **Allowing for uncertainty due to missing data in meta-analysis-Part 1 : Two-stage methods.** *Statistics in Medicine* 2008, **27**, pp. 711-727

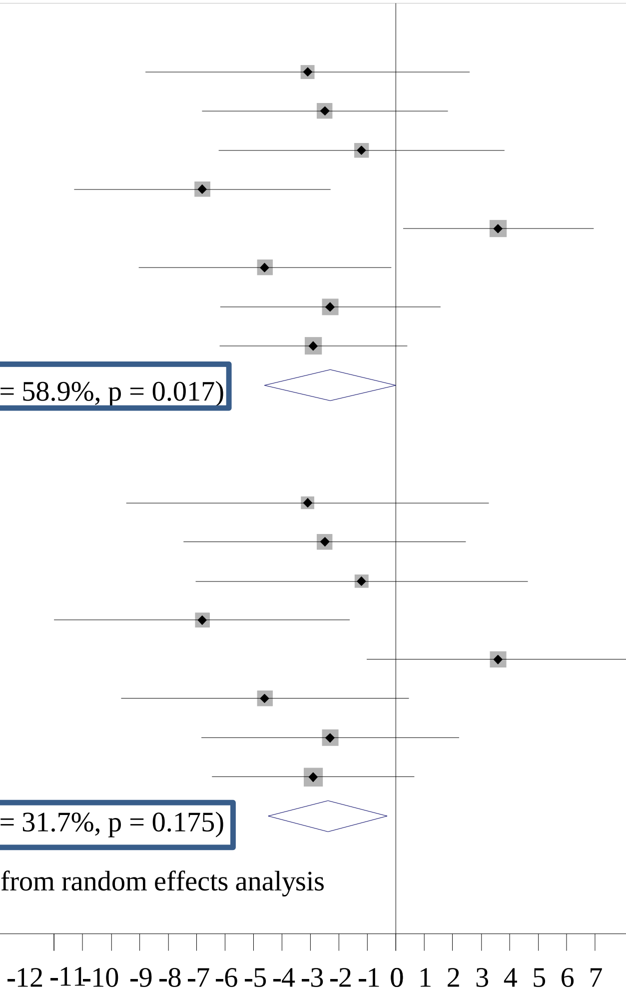
Mirtazapine vs placebo

ACA

IMDOM

$\mu_\lambda = 0, \sigma_\lambda = 2$

study	ES (95% CI)	% Weight
Primary analysis		
Claghorn 1995	-3.10 (-8.80, 2.60)	9.37
MIR 003-003	-2.50 (-6.81, 1.81)	12.37
MIR 003-008	-1.20 (-6.22, 3.82)	10.75
MIR 003-020	-6.80 (-11.30, -2.30)	11.91
MIR 003-021	3.60 (0.25, 6.95)	14.92
MIR 003-024	-4.60 (-9.04, -0.16)	12.07
MIR 84023a	-2.30 (-6.17, 1.57)	13.52
MIR 84023b	-2.90 (-6.19, 0.39)	15.09
Subtotal (I-squared = 58.9%, p = 0.017)	-2.31 (-4.62, 0.01)	100.00
.		
Secondary analysis		
Claghorn 1995	-3.10 (-9.46, 3.26)	8.50
MIR 003-003	-2.50 (-7.46, 2.46)	12.31
MIR 003-008	-1.20 (-7.04, 4.64)	9.72
MIR 003-020	-6.80 (-11.99, -1.61)	11.53
MIR 003-021	3.60 (-1.02, 8.22)	13.54
MIR 003-024	-4.60 (-9.65, 0.45)	11.98
MIR 84023a	-2.30 (-6.82, 2.22)	13.92
MIR 84023b	-2.90 (-6.46, 0.66)	18.50
Subtotal (I-squared = 31.7%, p = 0.175)	-2.39 (-4.48, -0.30)	100.00



NOTE: Weights are from random effects analysis

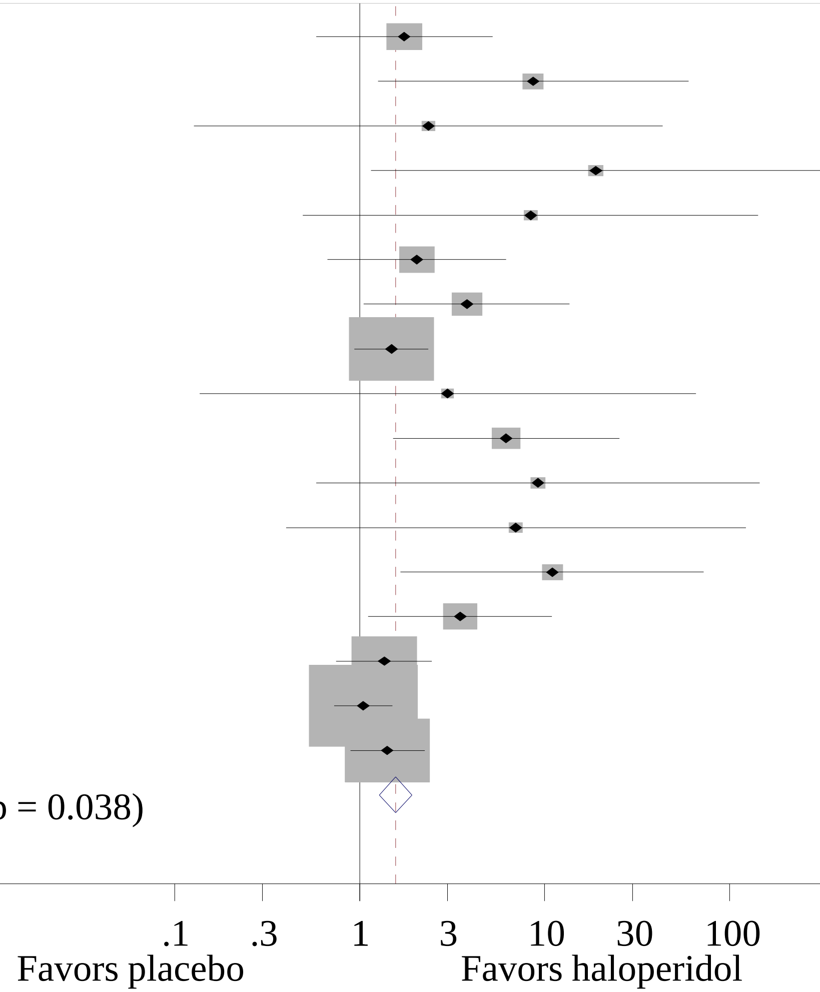
Most studies have zero missing values!

r: success **f**: failures **m**:missing	Haloperidol			Placebo		
	rh	**fh**	**mh**	**rp**	**fp**	**mp**
Arvanitis	25	25	2	18	[illegible]	0
Beasley	29	18	22	20	[illegible]	34
Bechelli	12	17	1	[illegible]	[illegible]	1
Borison	3	9	0	[illegible]	[illegible]	0
Chouinard	10	11	0	[illegible]	19	0
Durost	11	8	[illegible]	[illegible]	14	0
Garry	7	18	[illegible]	[illegible]	21	1
Howard	8	9	[illegible]	3	10	0
Marder	19	45	[illegible]	14	50	2
Nishikawa 82	1	[illegible]	[illegible]	0	10	0
Nishikawa 84	11	[illegible]	3	0	13	0
Reschke	20	[illegible]	0	2	9	0
Selman	[illegible]	[illegible]	11	7	4	18
Serafetinides	[illegible]	[illegible]	0	0	13	1
Simpson	[illegible]	14	0	0	7	1
Spencer	11	1	0	1	11	0
Vichaiya	9	20	1	0	29	1

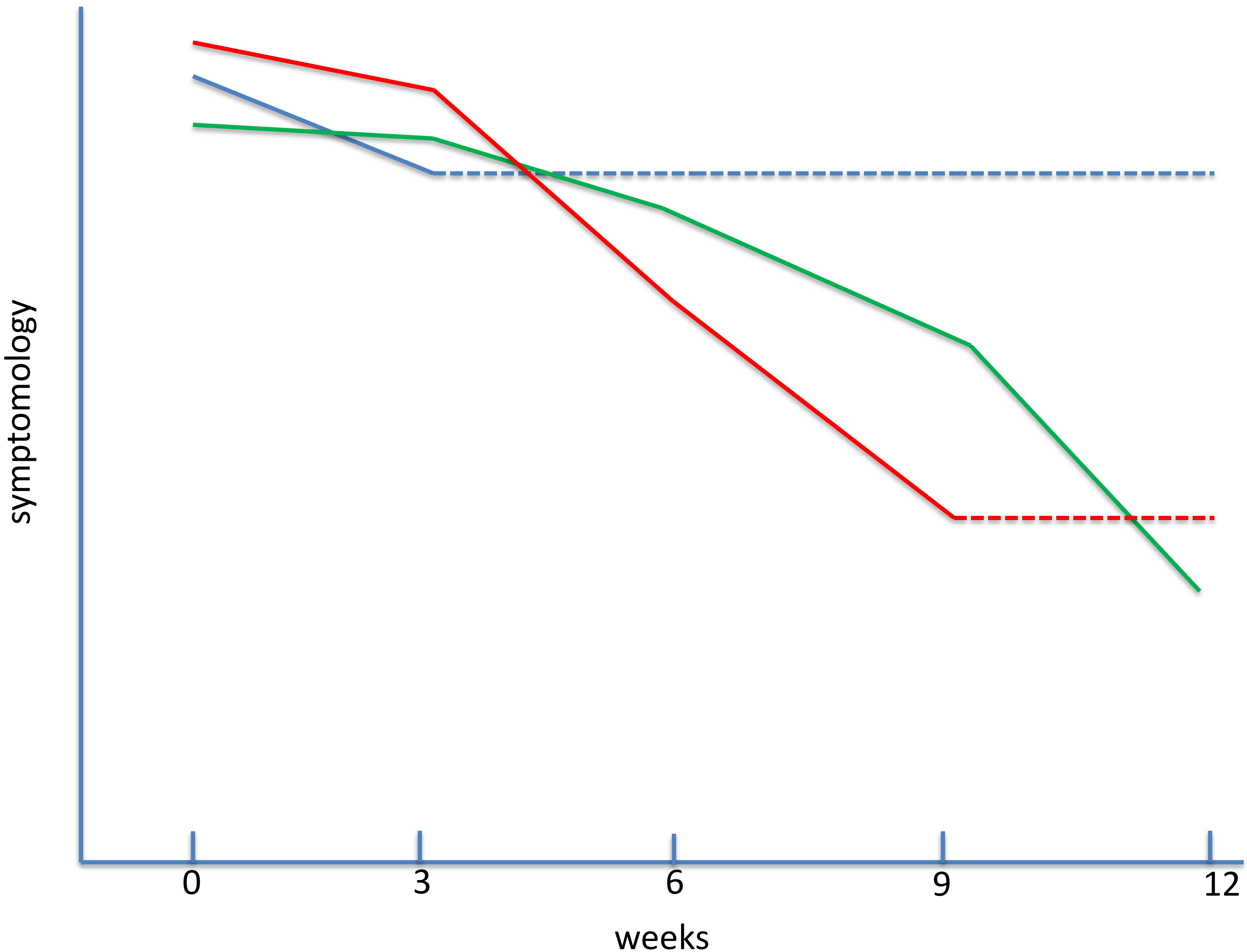
BECAUSE THEY USED A SINGLE IMPUTATION METHOD (LOCF)

Haloperidol vs placebo

author	RR (95% CI)	% Weight
Garry	1.75 (0.58, 5.24)	3.37
Durost	8.68 (1.26, 59.95)	1.09
Simpson	2.35 (0.13, 43.53)	0.48
Vichaiya	19.00 (1.16, 311.96)	0.52
Serafetinides	8.40 (0.50, 142.27)	0.51
Howard	2.04 (0.67, 6.21)	3.27
Reschke	3.79 (1.06, 13.60)	2.48
Selman	1.48 (0.94, 2.35)	19.11
Nishikawa_82	3.00 (0.14, 65.90)	0.42
Bechelli	6.21 (1.52, 25.35)	2.05
Nishikawa_84	9.20 (0.58, 145.76)	0.53
Borison	7.00 (0.40, 122.44)	0.49
Spencer	11.00 (1.67, 72.40)	1.14
Chouinard	3.49 (1.11, 10.95)	3.10
Marder	1.36 (0.75, 2.47)	11.37
Beasley	1.05 (0.73, 1.50)	**31.22**
Arvanitis	1.42 (0.89, 2.25)	18.86
Overall (I-squared = 41.4%, p = 0.038)	1.57 (1.28, 1.92)	100.00



missing rate 41%


symptomology
0
3
6
9
12
weeks

Conclusions

- Many popular statistical techniques to account for missing data do not take uncertainty of the imputed values into account and get smaller standard errors and confidence intervals
- we suggest models that can
  - account for the fact that the presence of missing data introduce uncertainty in the study estimates
  - naturally downweight studies with lots of missing data
  - can model MAR or departures from MAR
- we need priors for IM parameters or conduct a sensitivity analysis
- We have extended the approach to network meta-analysis
- metamiss command in STATA (Ian White & Julian Higgins); metamiss2 command in STATA (Anna Chaimani and Ian White, forthcoming)
- We extend the method to account for data that have been imputed using single imputation techniques (e.g. LOCF)

References

- Higgins JPT, White IR, Wood AM: Imputation methods for missing outcome data in meta-analysis of clinical trials. *Clinical Trials* 2008; 5, pp. 225-239
- Higgins JPT, Altman DG, Sterne JAC (editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
- Higgins JPT, Deeks JJ, Altman DG (editors). Chapter 16: Special topics in statistics. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
- Joy CB, Adams,CE, Lawrie,SM. Haloperidol versus placebo for schizophrenia. Cochrane Database Syst Rev 2006, 4. CD003082
- Mavridis D, Chaimani A, Efthimiou O, Leucht S, Salanti G.:Addressing missing outcome data in meta-analysis. Statistics in Practice. *Evidence Based Mental Health* 2014; 17, pp 85-89.
- Mavridis D., White I., Higgins J., Cipriani A., Salanti G Allowing for uncertainty due to missing continuous missing outcome data in pairwise and network meta-analysis. *Statistics in Medicine* 2014, *34, 721–741*
- White IR, Higgins JPT, Wood AM: Allowing for uncertainty due to missing data in meta-analysis- Part 1 : Two-stage methods. *Statistics in Medicine* 2008, 27, pp. 711-727
- White IR, Welton NJ, Wood AM, Ades AE, Higgins JPT: Allowing for uncertainty due to missing data in meta-analysis -Part 2 : Hierarchical models. *Statistics in Medicine* 2008; 27, pp. 728-745
- White IR, Higgins JPT. Meta-analysis with missing data. Stata J. 9(1):57–69.

Implementation in Stata

The `metamiss` command

`metamiss rE fE mE rC fC mC, imputation_method [options]`

`aca`	performs an available cases analysis
`ica0`	imputes missing values as failures
`ica1`	imputes missing values as successes
`icab`	performs a best-case analysis
`icaw`	performs a worst-case analysis
`icape`	imputes missing values by using the observed probability in the experimental group
`icapc`	imputes missing values by using the observed probability in the control group
`icap`	imputes missing values by using the observed probability within groups
`icaimor`	imputes missing values by using the IMORs specified by `imor()` or `logimor()` and `sdlogimor()` within groups

The `metamiss2` command

```
metamiss2 input_variables, IPM_definition [options]
```

`impmean()`	defines the mean(s) for the IMP parameter(s)
`impsd()`	defines the standard deviation(s) for the IMP parameters(s)
`impcorr()`	defines the correlation of the IMP parameters across groups
`compare()`	runs simultaneously two different analyses with different assumptions about the IMP parameter(s)
`sensitivity`	runs a sensitivity analysis on a range of different standard deviations for the IMP parameter

- Can be installed by typing (requires Stata 13 or later version):

```
net install metamiss2, from(http://www.mtm.uoi.gr)
```

Allowing for informative missingness in aggregate data meta-analysis with continuous or binary outcomes: 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. However, addressing missing data is often a challenging issue since it requires untestable assumptions. The impact of missing outcome data on the meta-analysis 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 IMPs cannot be estimated from the observed data but they can be specified, with associated uncertainty, using evidence external to the meta-analysis, such as expert

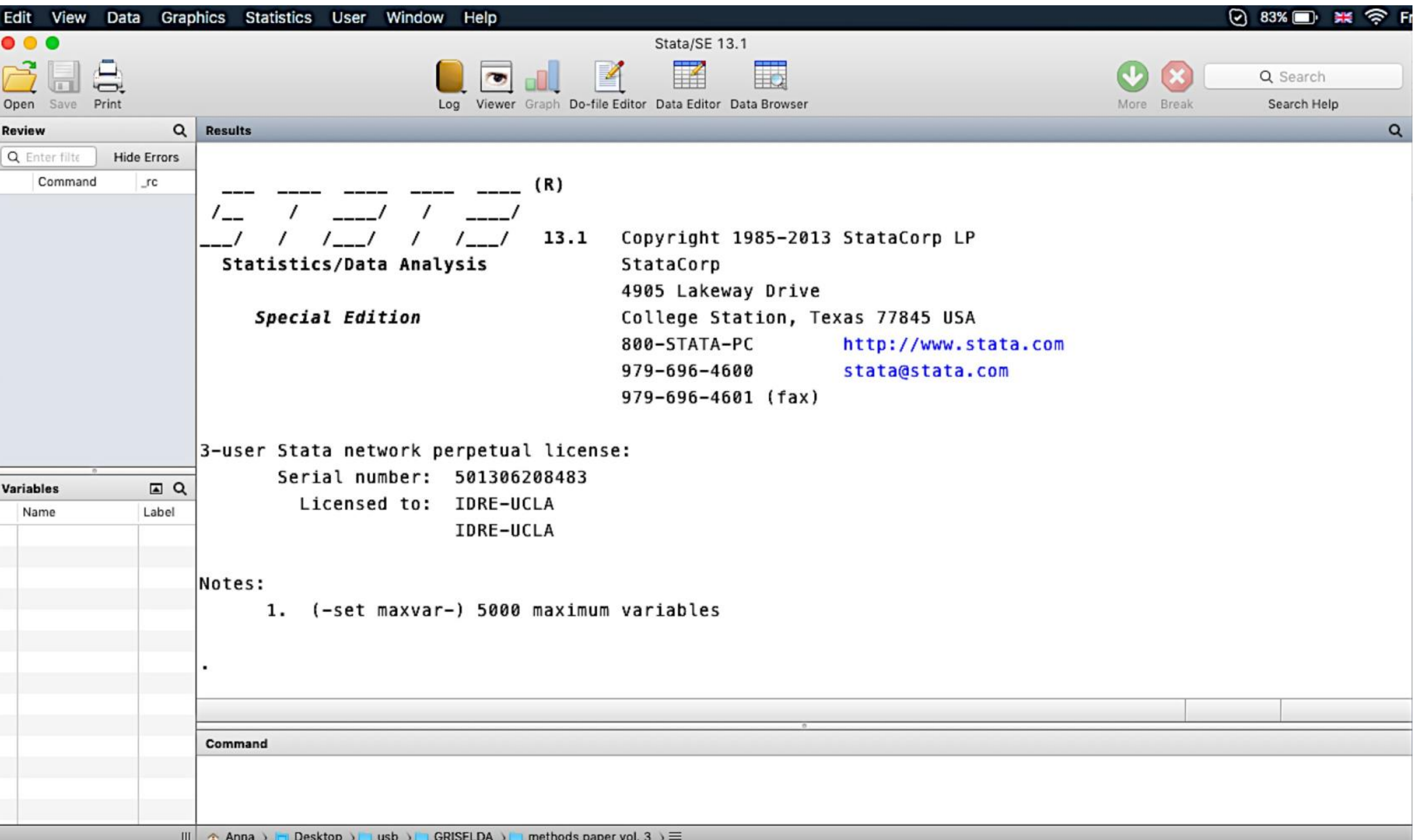
Installation of the Stata commands

If you have not already done so…

```
*******************************************************************************
*       Allowing for uncertainty due to missing outcome data in meta-analysis   *
*                         Anna Chaimani & Dimitris Mavridis                     *
*                         18 June 2018, Cochrane SMG webinar                    *
*******************************************************************************

*** 1.Installing the Stata commands ***

//if you have not already done so, please install the required Stata commands
cap ssc install metan
cap ssc install metamiss
net install metamiss2, replace from(http://www.mtm.uoi.gr)

*** 2.Pairwise meta-analysis - binary data ***
```

Pairwise meta-analysis – binary data

✓ Load the dataset “haloperidol.dta”



Pairwise meta-analysis – binary data

✓ Look at the dataset

missing data workshop commands Vienna 2015.do

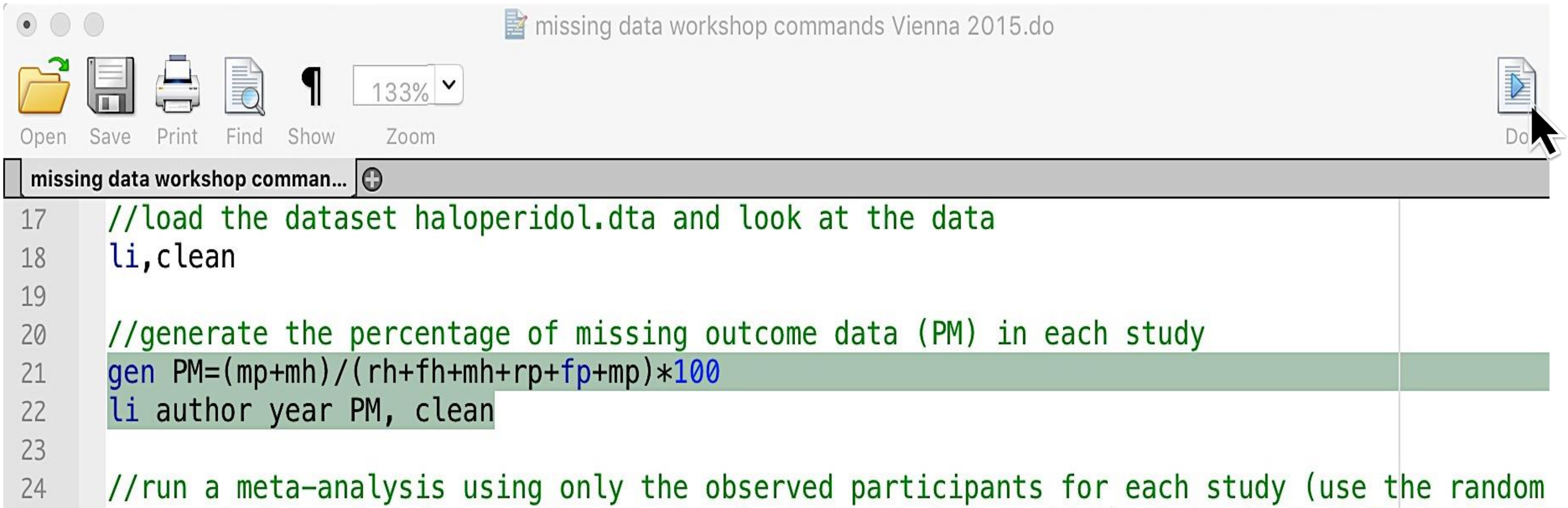
```
net install metamiss2, replace from(http://www.mtm.uoi.gr)

*** 2.Pairwise meta-analysis - binary data ***

//load the dataset haloperidol.dta and look at the data
li,clean

//generate the percentage of missing outcome data (PM) in each study
gen PM=(mp+mh)/(rh+fh+mh+rp+fp+mp)*100
li author year PM, clean

```

Pairwise meta-analysis – binary data

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

Pairwise meta-analysis – binary data

✓ Find the percentage of missing outcome data in each study



Pairwise meta-analysis – binary data

✓ Find the percentage of missing outcome data in each study

author	year	PM
Arvanitis	1997	1.941748
Beasley	1996	40.875910
Bechelli	1983	3.278688
Borison	1992	0
Chouinard	1993	0
Durost	1964	0
Garry	1962	3.846154
Howard	1974	0
Marder	1994	3.030303
Nishikawa_82	1982	0
Nishikawa_84	1984	6.000000
Reschke	1974	0
Selman	1976	50.000000
Serafetinides	1972	3.571429
Simpson	1967	4.166667
Spencer	1992	0
Vichaiya	1971	3.333333

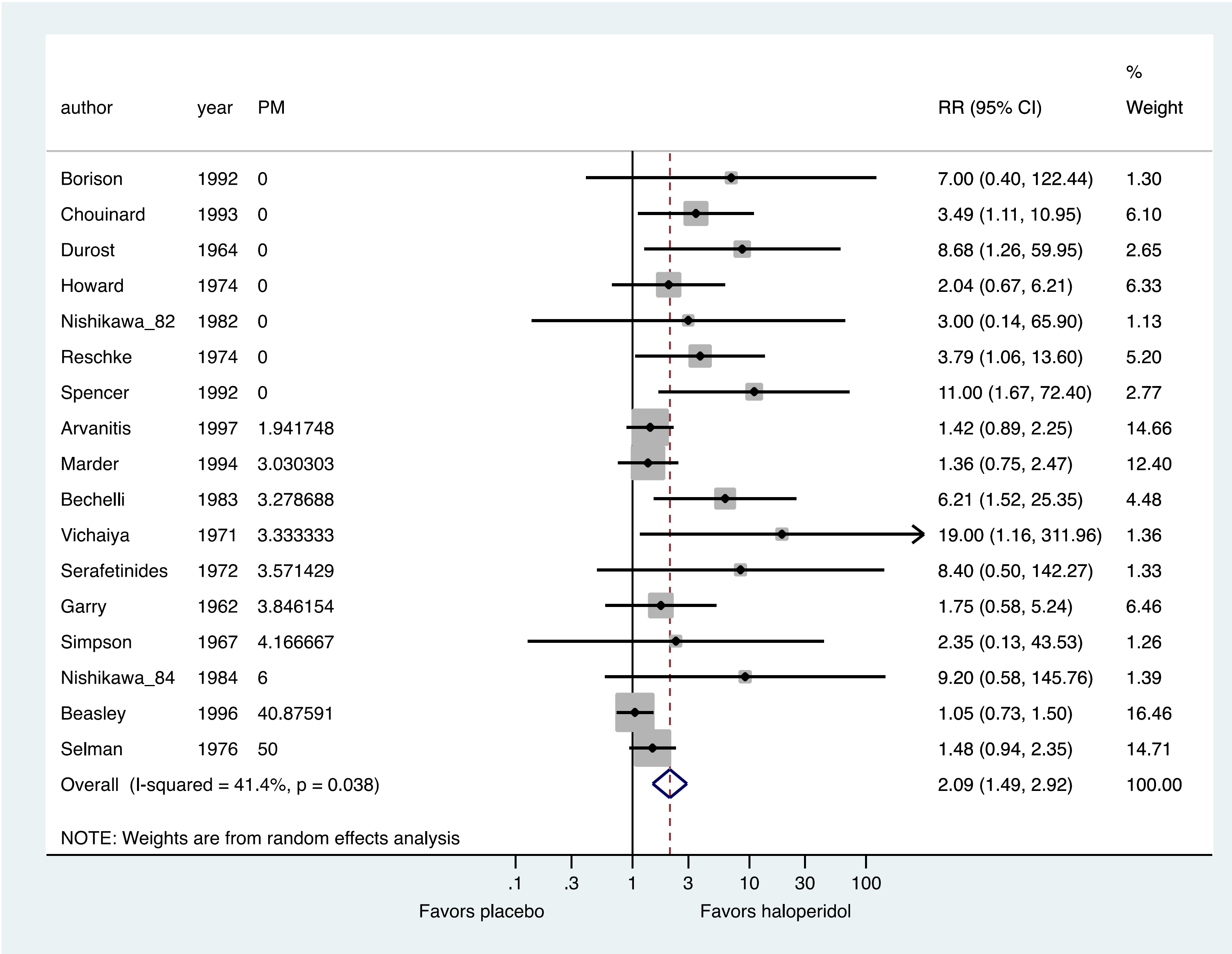
Pairwise meta-analysis – binary data

✓ Run a (random-effects meta-analysis) using only the observed participants (ACA)

missing data workshop commands Vienna 2015.do

```
//load the dataset haloperidol.dta and look at the data
li,clean

//generate the percentage of missing outcome data (PM) in each study
gen PM=(mp+mh)/(rh+fh+mh+rp+fp+mp)*100
li author year PM, clean

/*run a meta-analysis using only the observed participants for each study
(use the random effects model); that is the available-case analysis (ACA)*/
metan rh fh rp fp, favours(Favors placebo # Favors haloperidol) ///
xlabel(0.1, 0.3, 1, 3, 10, 30, 100) sortby(PM) lcols(author year PM) randomi texts(130)
```

Pairwise meta-analysis – binary data



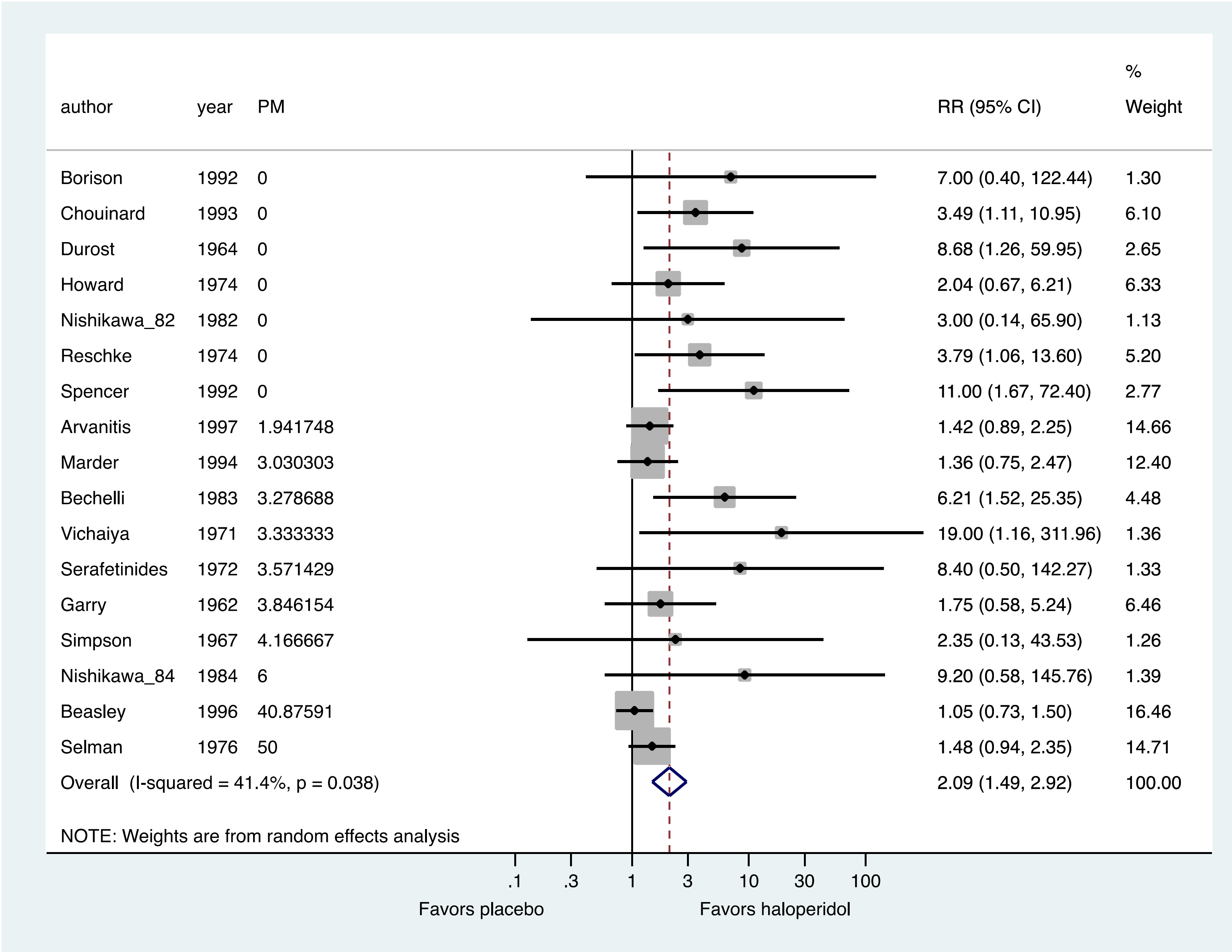
Pairwise meta-analysis – binary data

✓ Run the same analysis using `metamiss2` instead of `metan`

missing data workshop commands Vienna 2015.do

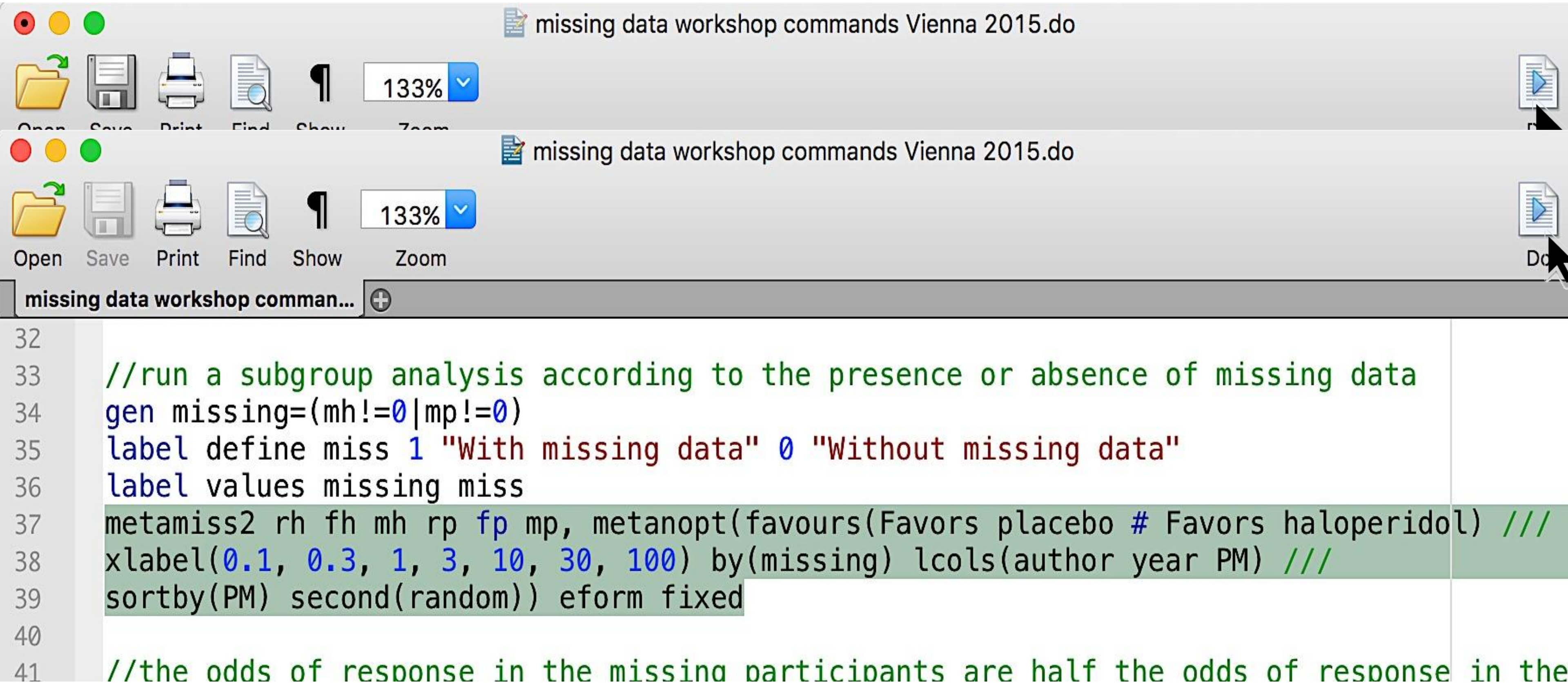
```
metan rh fh rp fp, favours(Favors placebo # Favors haloperidol) ///
xlabel(0.1, 0.3, 1, 3, 10, 30, 100) sortby(PM) lcols(author year PM) randomi texts(130)

//the above command is equivalent with the following
metamiss2 rh fh mh rp fp mp, metanopt(favours(Favors placebo # Favors haloperidol) ///
xlabel(0.1, 0.3, 1, 3, 10, 30, 100) lcols(author year PM) sortby(PM) texts(130)) eform

//run a subgroup analysis according to the presence or absence of missing data
```

Pairwise meta-analysis – binary data

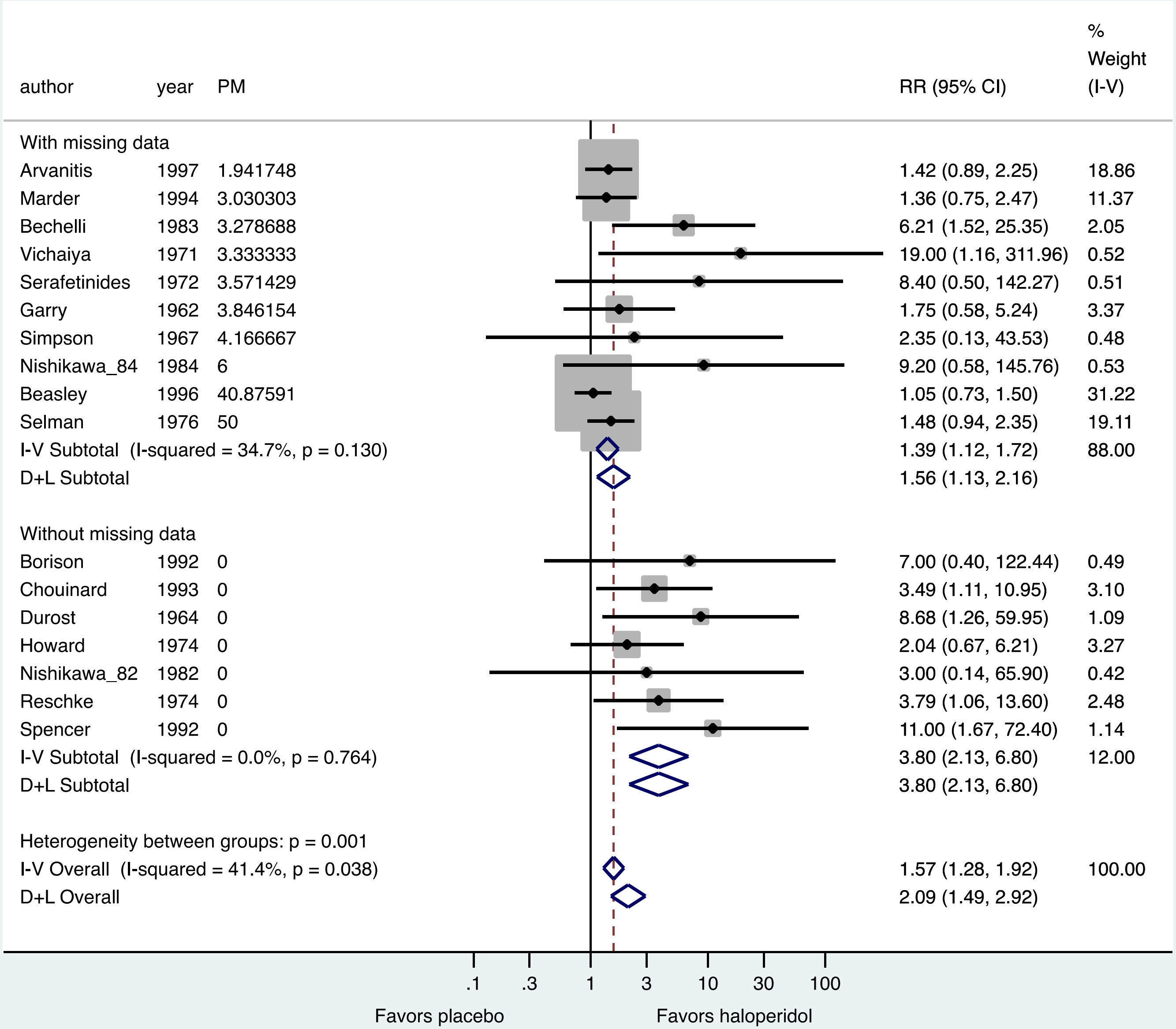


Pairwise meta-analysis – binary data

- ✓ Create a variable indicating the presence or absence of missing data for each study
- ✓ Run a subgroup analysis using this variable



Pairwise meta-analysis – binary data



Pairwise meta-analysis – binary data

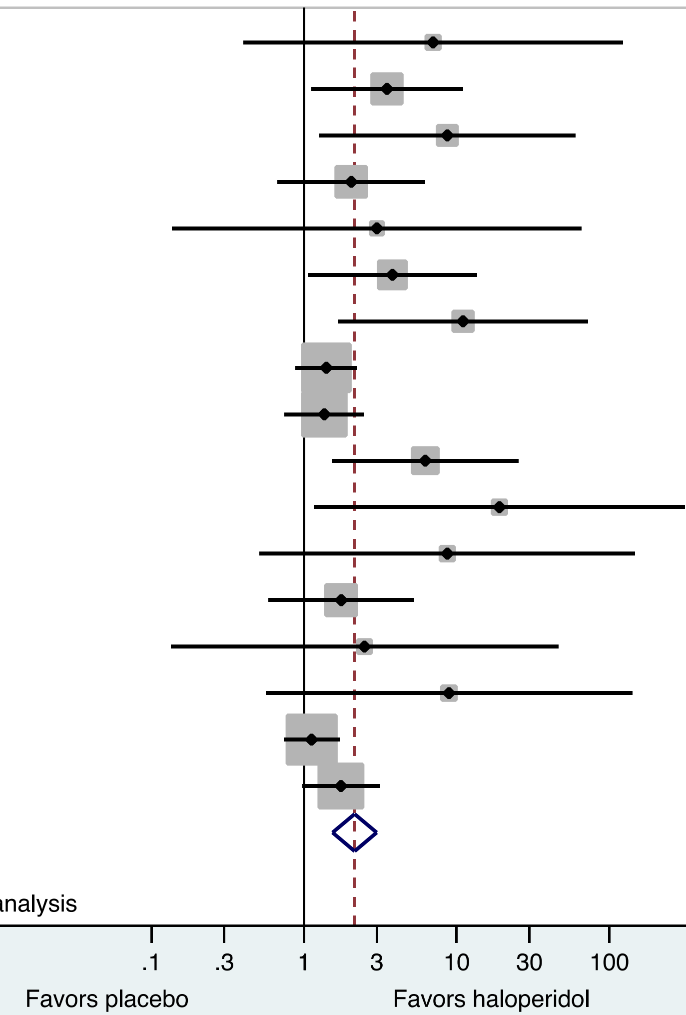
✓ Assume that the odds of response in the missing group are half the odds of response in the observed data for both arms

missing data workshop commands Vienna 2015.do

```
xlabel(0.1, 0.3, 1, 3, 10, 30, 100) by(missing) lcols(author year PM) ///
sortby(PM) second(random)) eform fixed

/*the odds of response in the missing participants are half the odds of response
in the observed data for both arms*/
di log(0.5)
metamiss2 rh fh mh rp fp mp, impmean(-0.7) eform ///
metanopt(favours(Favors placebo # Favors haloperidol) ///
xlabel(0.1, 0.3, 1, 3, 10, 30, 100) lcols(author year PM) sortby(PM) texts(130))

//the same results are obtained when we run the same analysis with metamiss
```

Pairwise meta-analysis – binary data

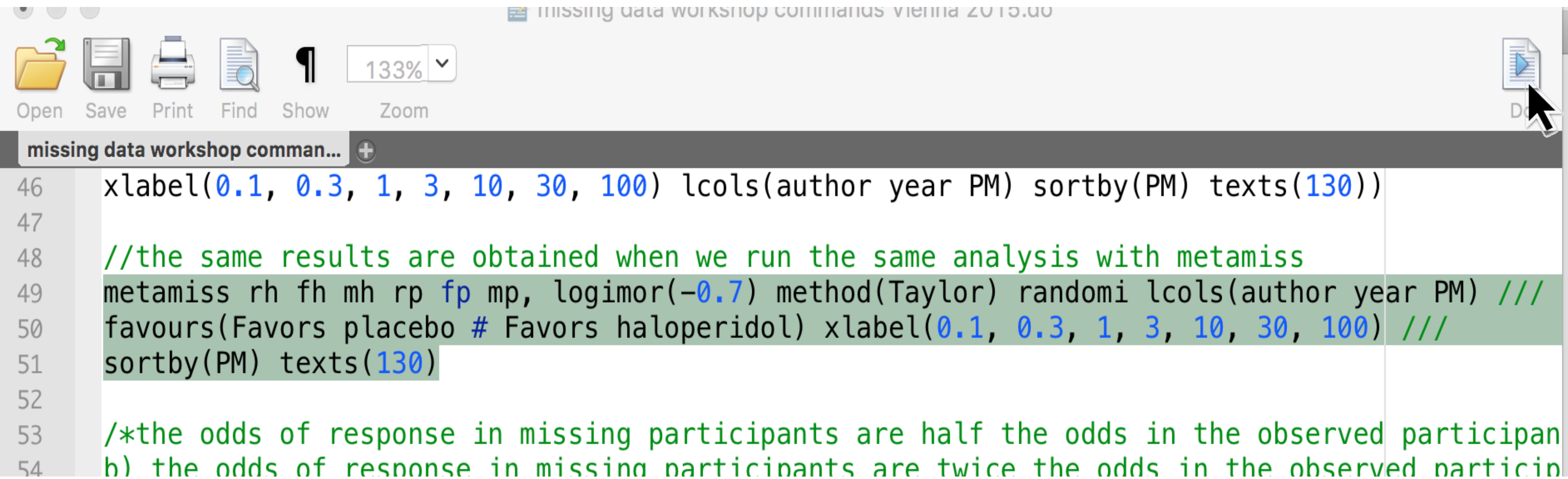
author	year	PM	RR (95% CI)	% Weight
Borison	1992	0	7.00 (0.40, 122.44)	1.28
Chouinard	1993	0	3.49 (1.11, 10.95)	6.15
Durost	1964	0	8.68 (1.26, 59.95)	2.62
Howard	1974	0	2.04 (0.67, 6.21)	6.39
Nishikawa_82	1982	0	3.00 (0.14, 65.90)	1.10
Reschke	1974	0	3.79 (1.06, 13.60)	5.21
Spencer	1992	0	11.00 (1.67, 72.40)	2.74
Arvanitis	1997	1.941748	1.40 (0.88, 2.23)	15.56
Marder	1994	3.030303	1.36 (0.75, 2.47)	12.97
Bechelli	1983	3.278688	6.23 (1.52, 25.49)	4.47
Vichaiya	1971	3.333333	19.06 (1.16, 313.20)	1.33
Serafetinides	1972	3.571429	8.69 (0.51, 147.47)	1.30
Garry	1962	3.846154	1.75 (0.58, 5.27)	6.50
Simpson	1967	4.166667	2.49 (0.13, 46.33)	1.23
Nishikawa_84	1984	6	8.91 (0.56, 141.24)	1.36
Beasley	1996	40.87591	1.12 (0.74, 1.70)	16.53
Selman	1976	50	1.75 (0.97, 3.13)	13.25
Overall (I-squared = 35.0%, p = 0.077)			2.14 (1.54, 2.99)	100.00



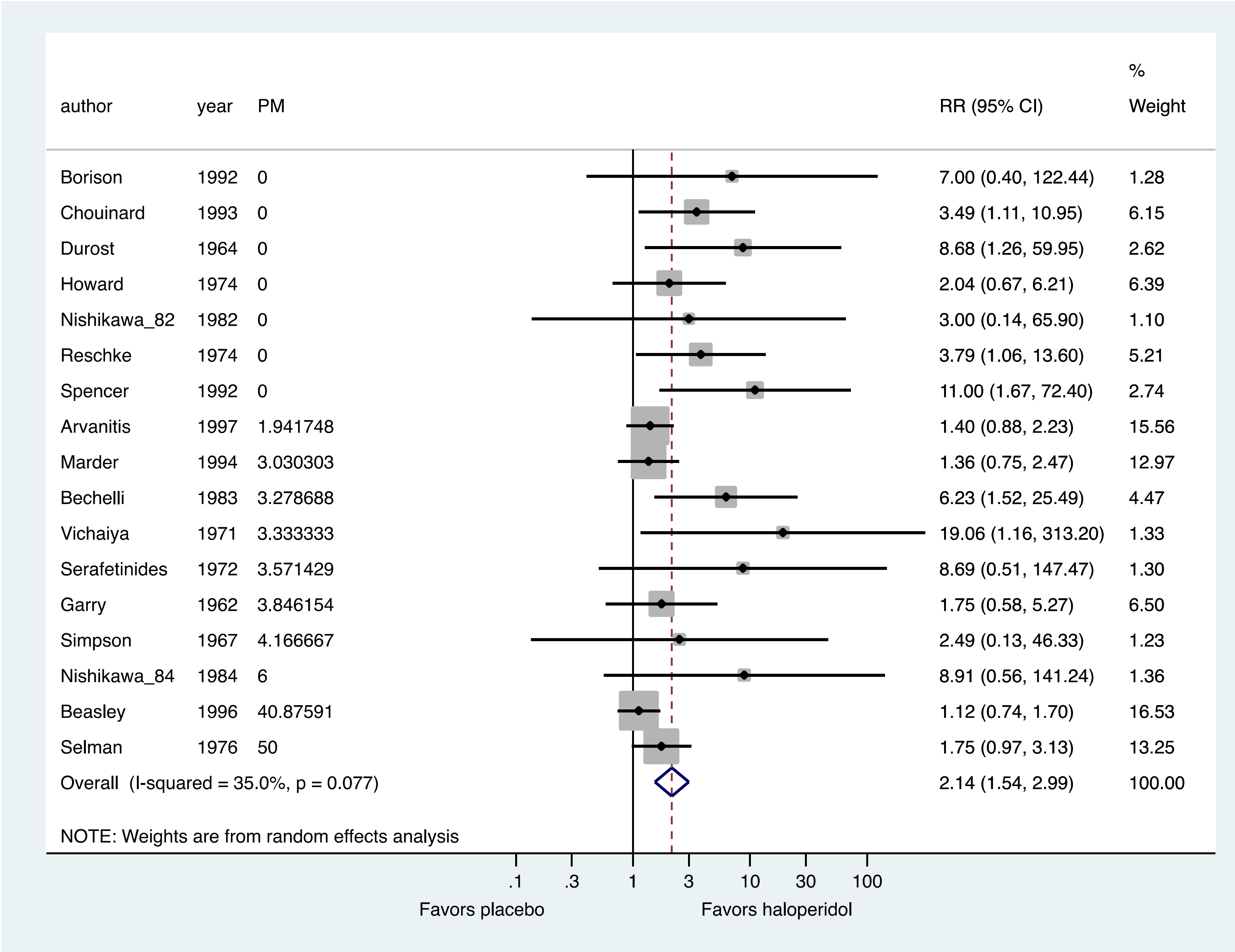
NOTE: Weights are from random effects analysis

Pairwise meta-analysis – binary data

✓ Run the same analysis using `metamiss` instead of `metamiss2`

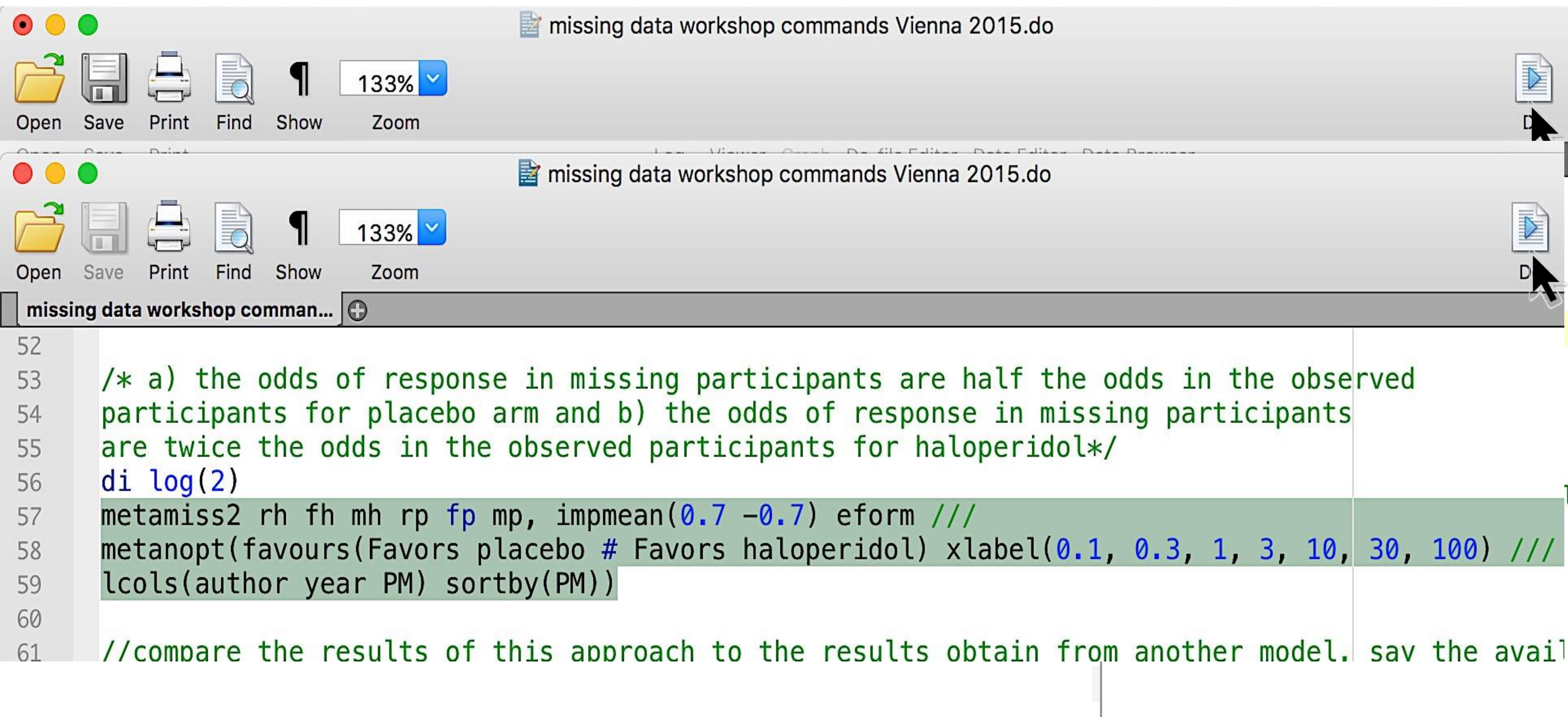


Pairwise meta-analysis – binary data

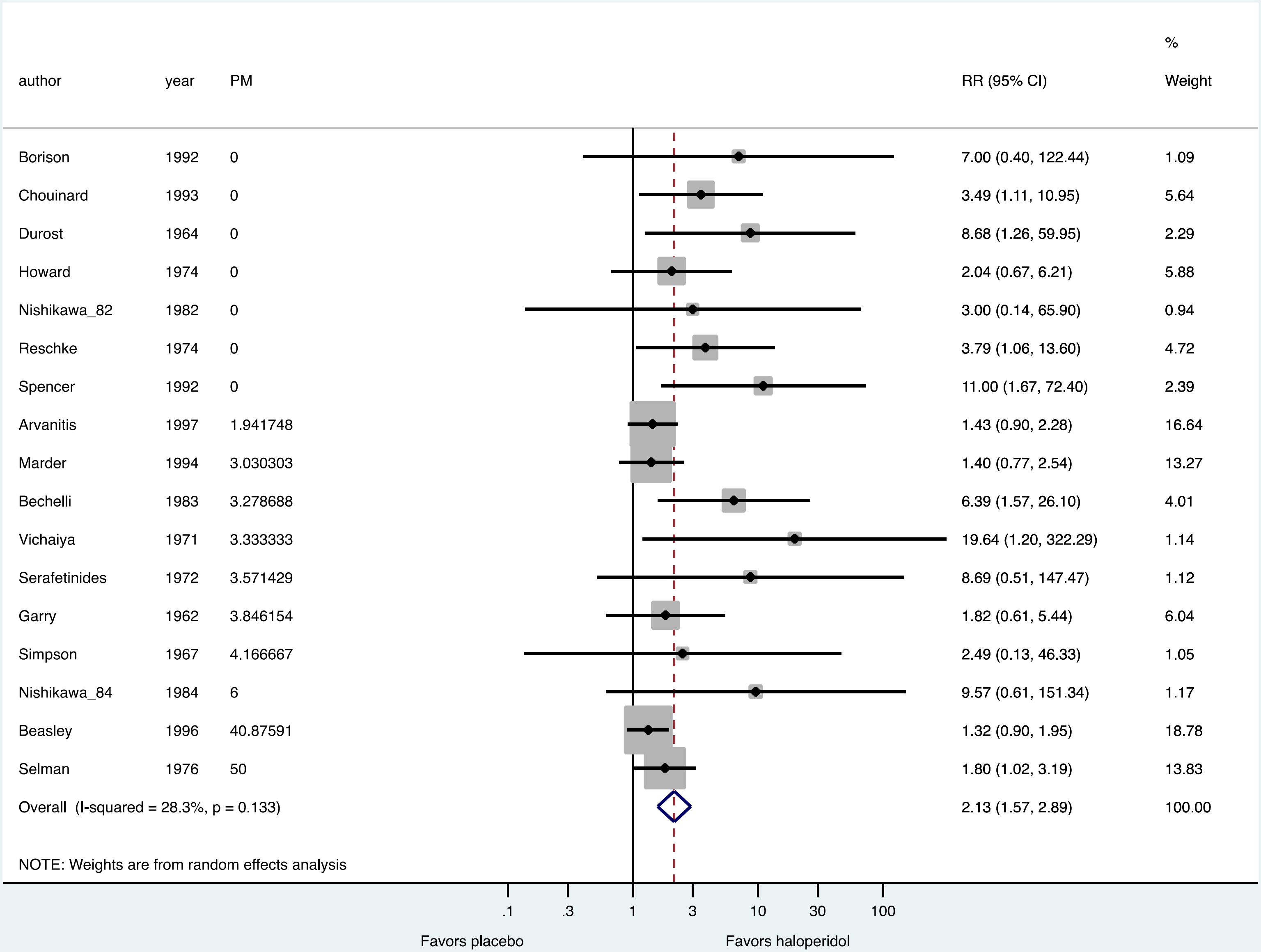


Pairwise meta-analysis – binary data

- ✓ a) the odds of response in missing participants are half the odds in the observed participants for placebo arm and b) the odds of response in missing participants are twice the odds in the observed participants for haloperidol

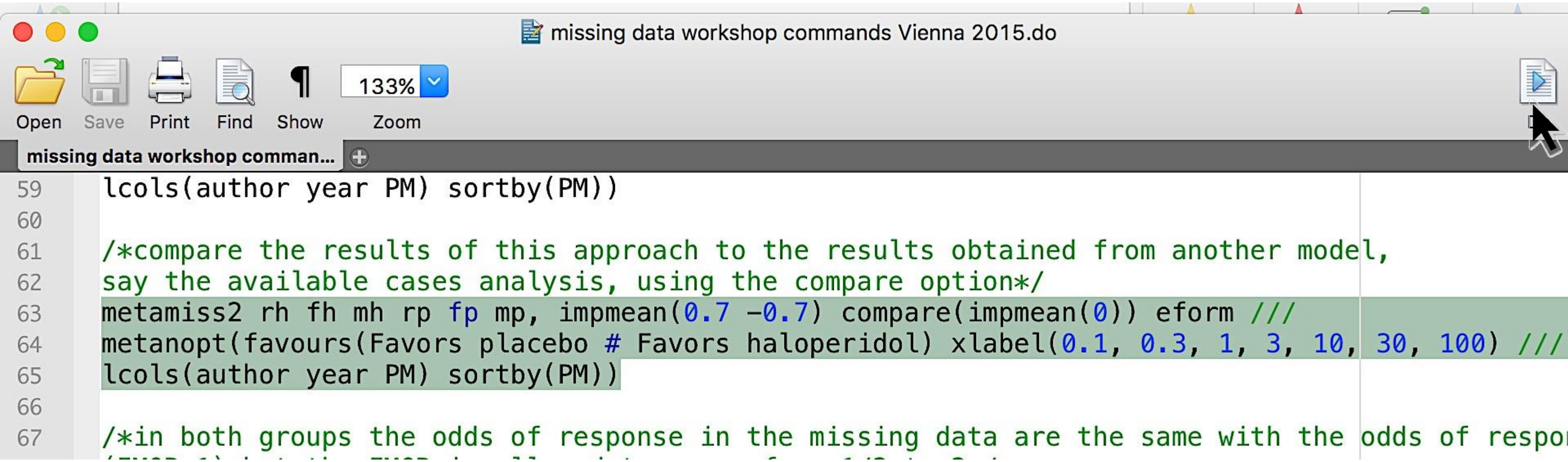


Pairwise meta-analysis – binary data

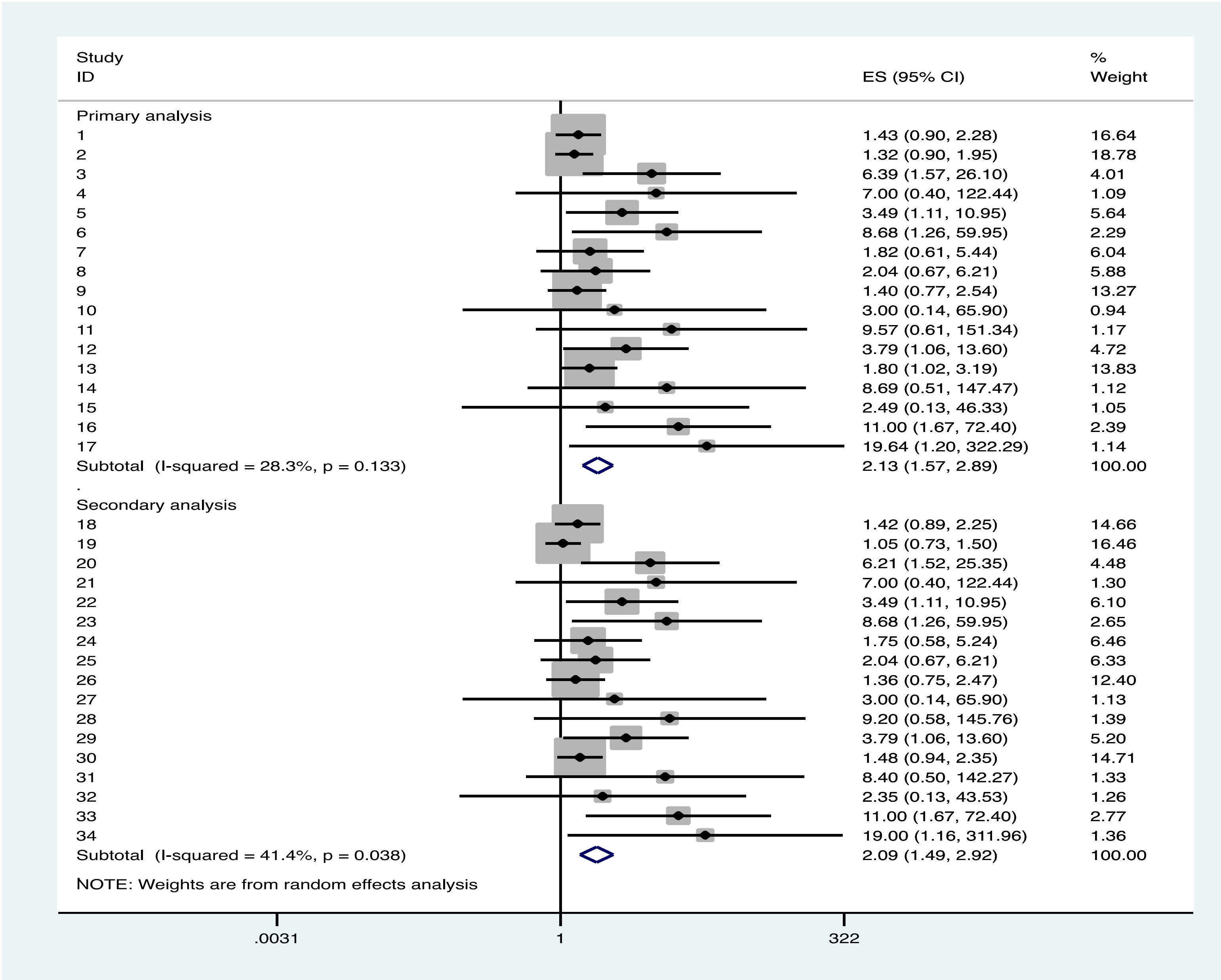


Pairwise meta-analysis – binary data

✓ Compare the results of this analysis to the results obtained from another model, say the available cases analysis

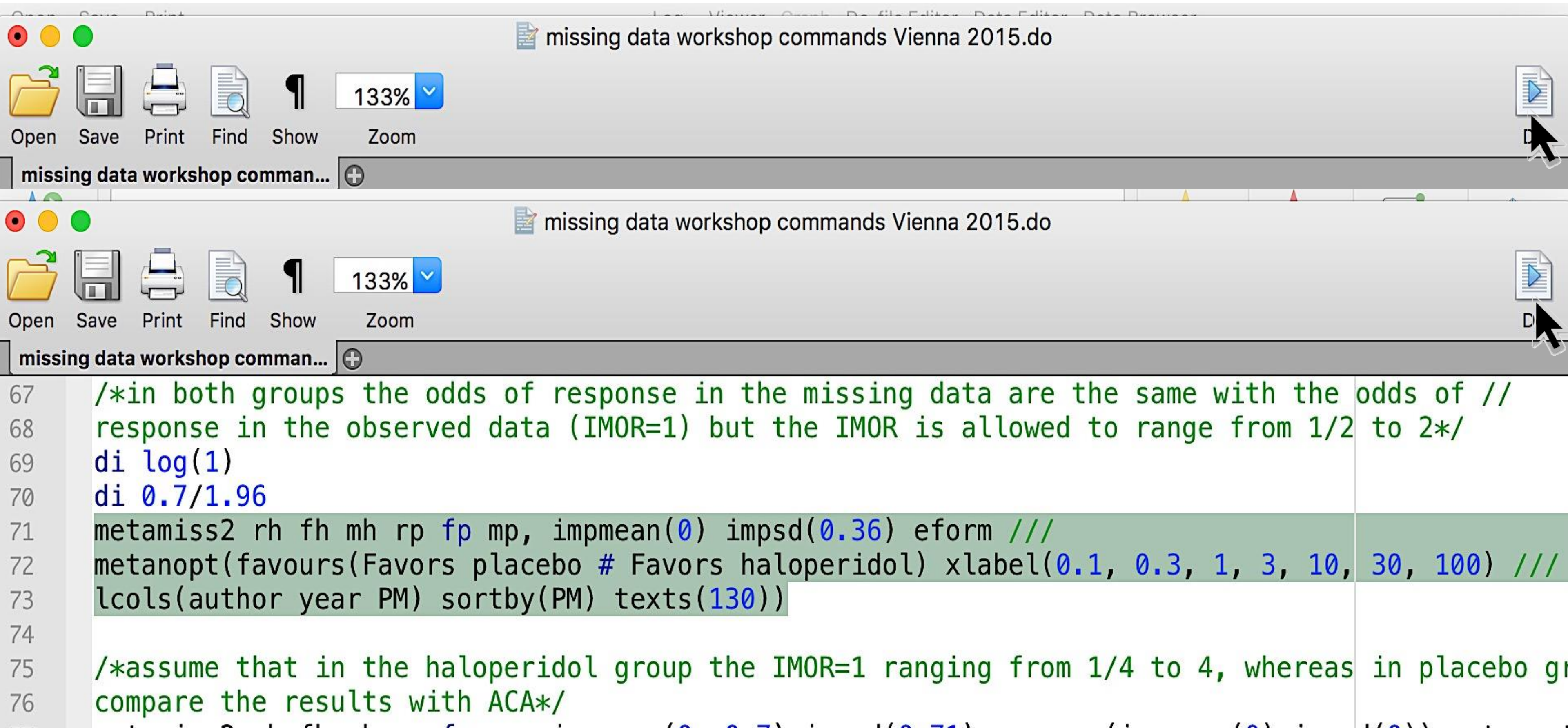


Pairwise meta-analysis – binary data

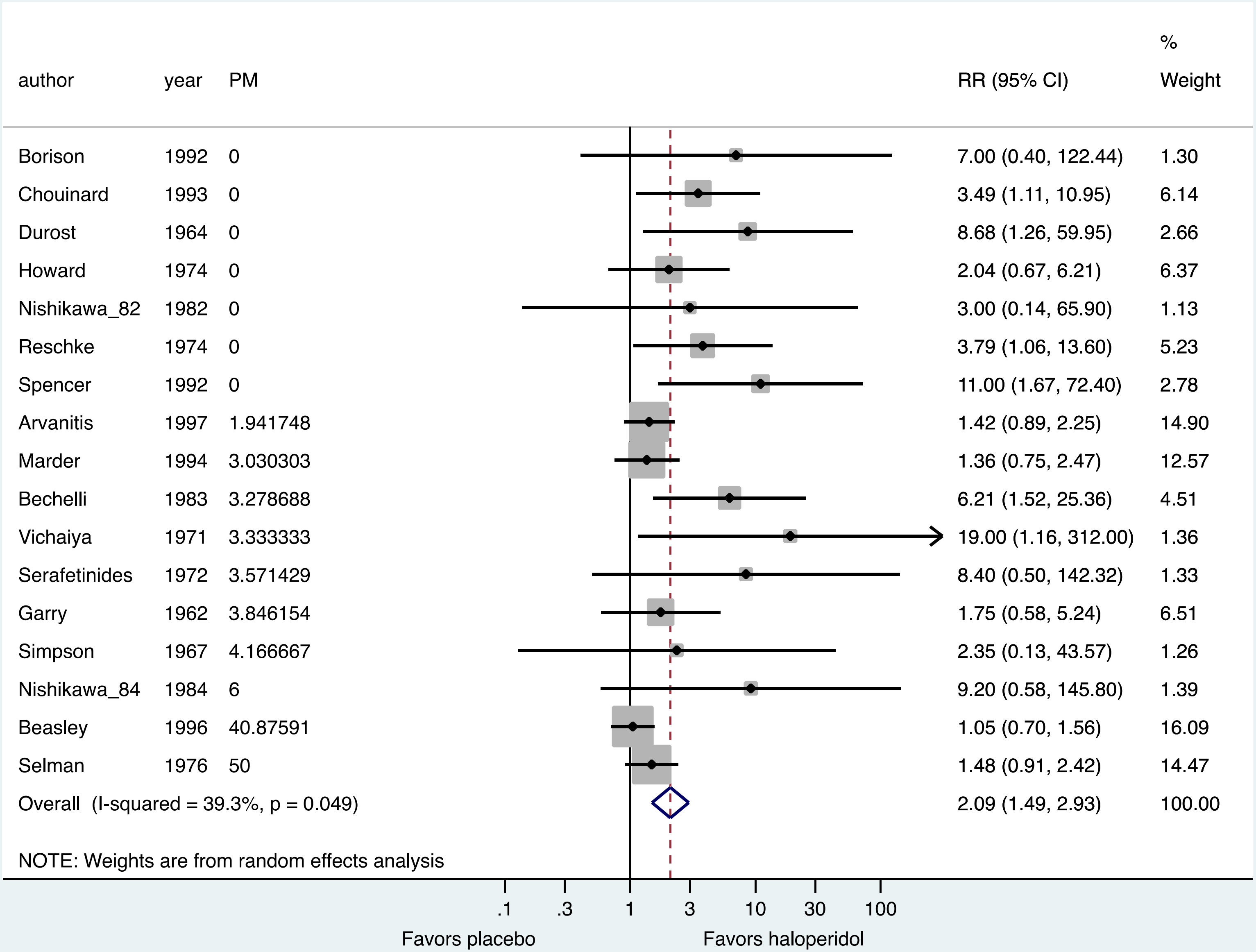


Pairwise meta-analysis – binary data

✓ Assume ACA in both groups (IMOR=1) but the IMOR is allowed to range from 1/2 to 2

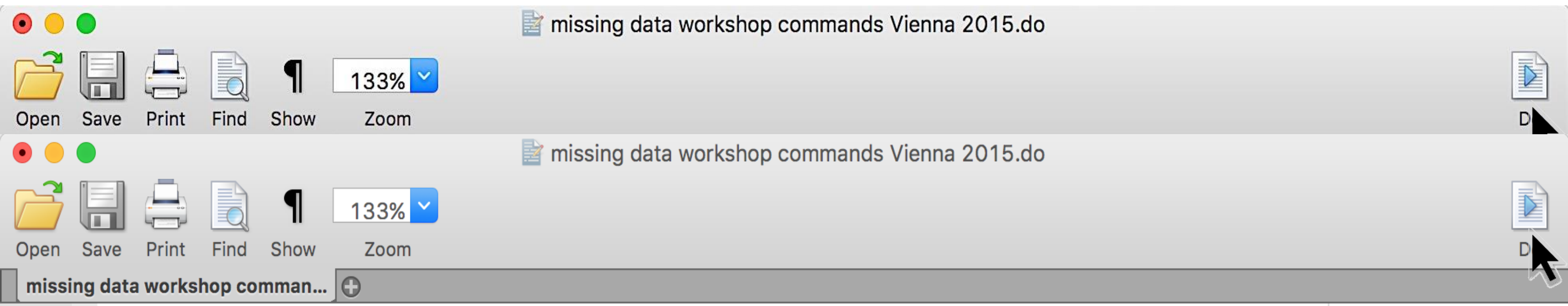


Pairwise meta-analysis – binary data



Pairwise meta-analysis – binary data

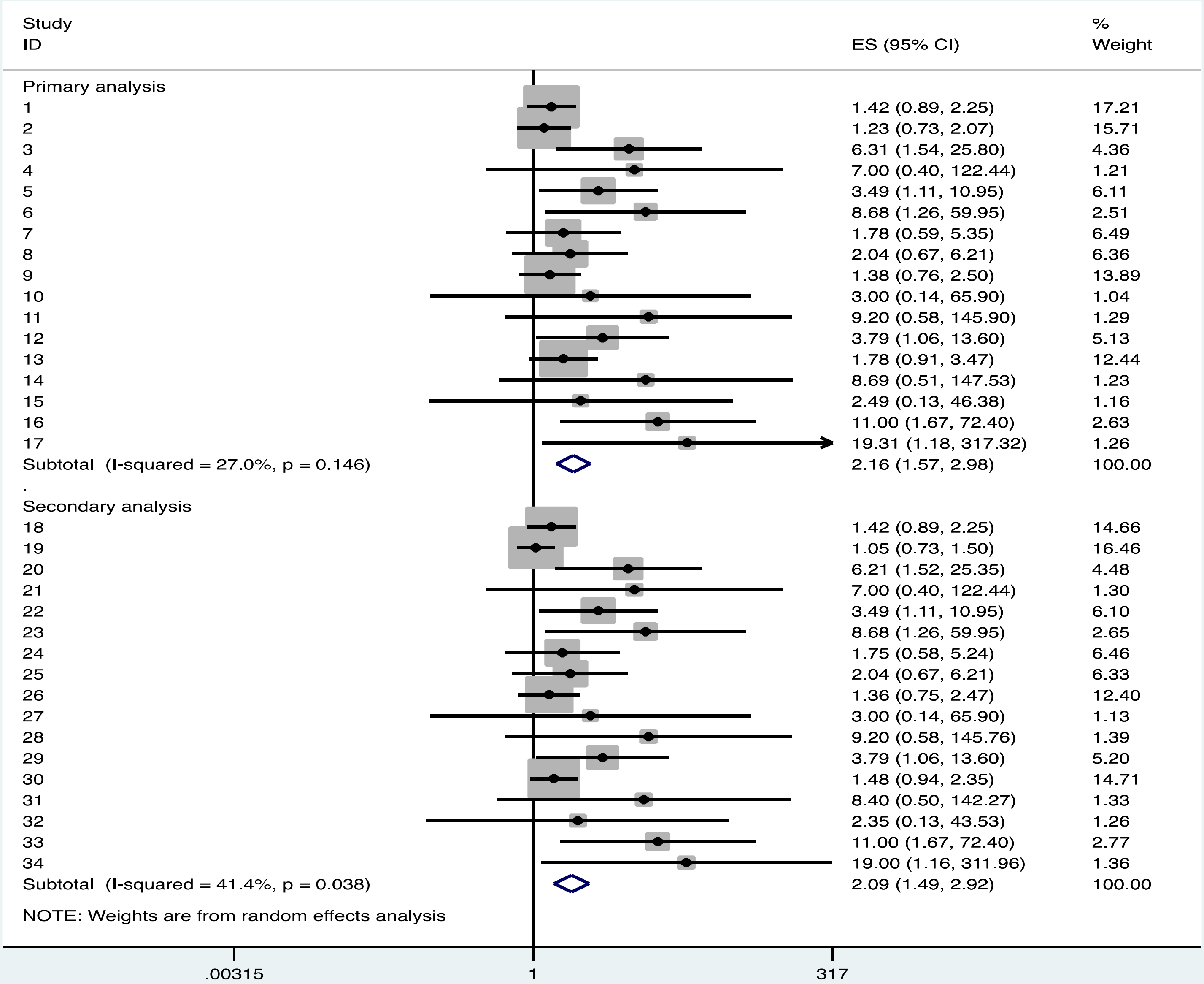
- ✓ Assume that in the haloperidol group the IMOR=1 ranging from 1/4 to 4, whereas in placebo group IMOR=1/2 ranging from 1/8 to 2, compare the results with ACA




```
/*assume that in the haloperidol group the IMOR=1 ranging from 1/4 to 4,
whereas in placebo group IMOR=1/2 ranging from 1/8 to 2, compare the results with ACA*/
di log(4)/1.96
di (0.7+0.7)/1.96
metamiss2 rh fh mh rp fp mp, impmean(0 -0.7) impsd(0.71) compare(impmean(0) impsd(0)) eform ///
metanopt(favours(Favors placebo # Favors haloperidol) xlabel(0.1, 0.3, 1, 3, 10, 30, 100) ///
lcols(author year PM) sortby(PM))

//run a sensitivity analysis showing how the uncertainty in IMOR impacts the summary effect wher
```

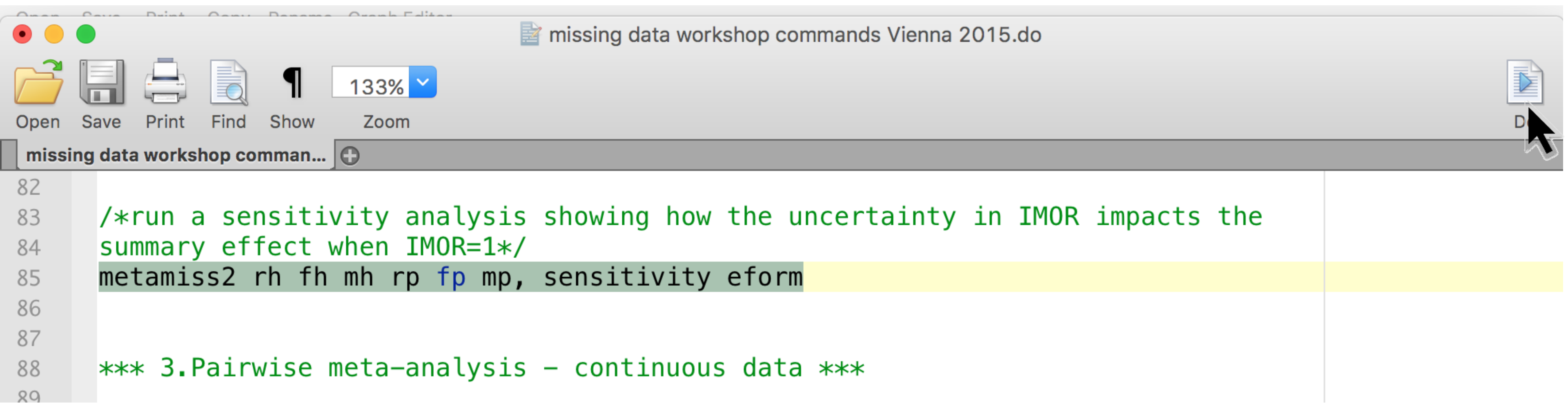
```
.
end of do-file
```

Pairwise meta-analysis – binary data

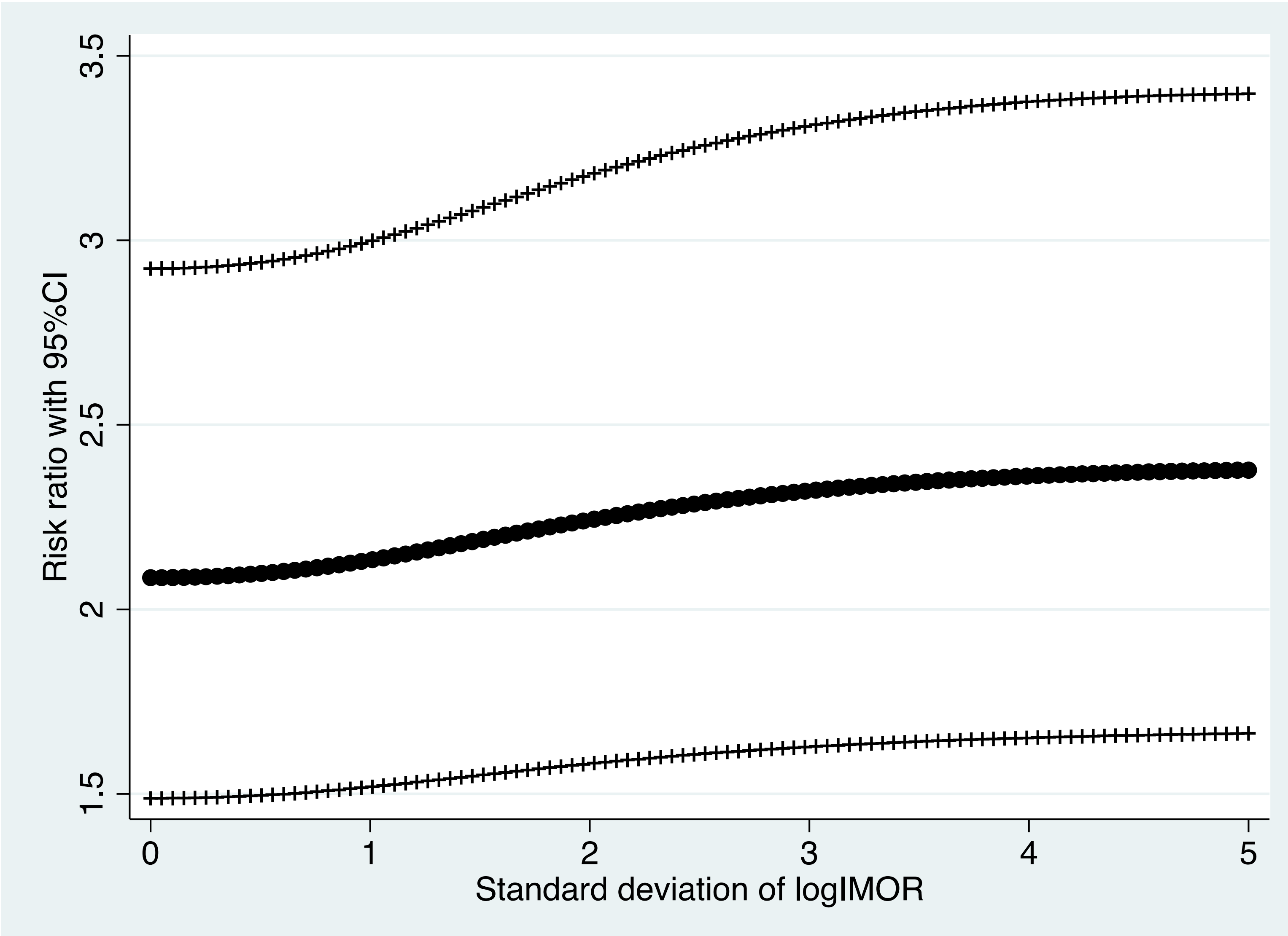


Pairwise meta-analysis – binary data

- ✓ Run a sensitivity analysis showing how the uncertainty in IMOR impacts the summary effect when IMOR=1

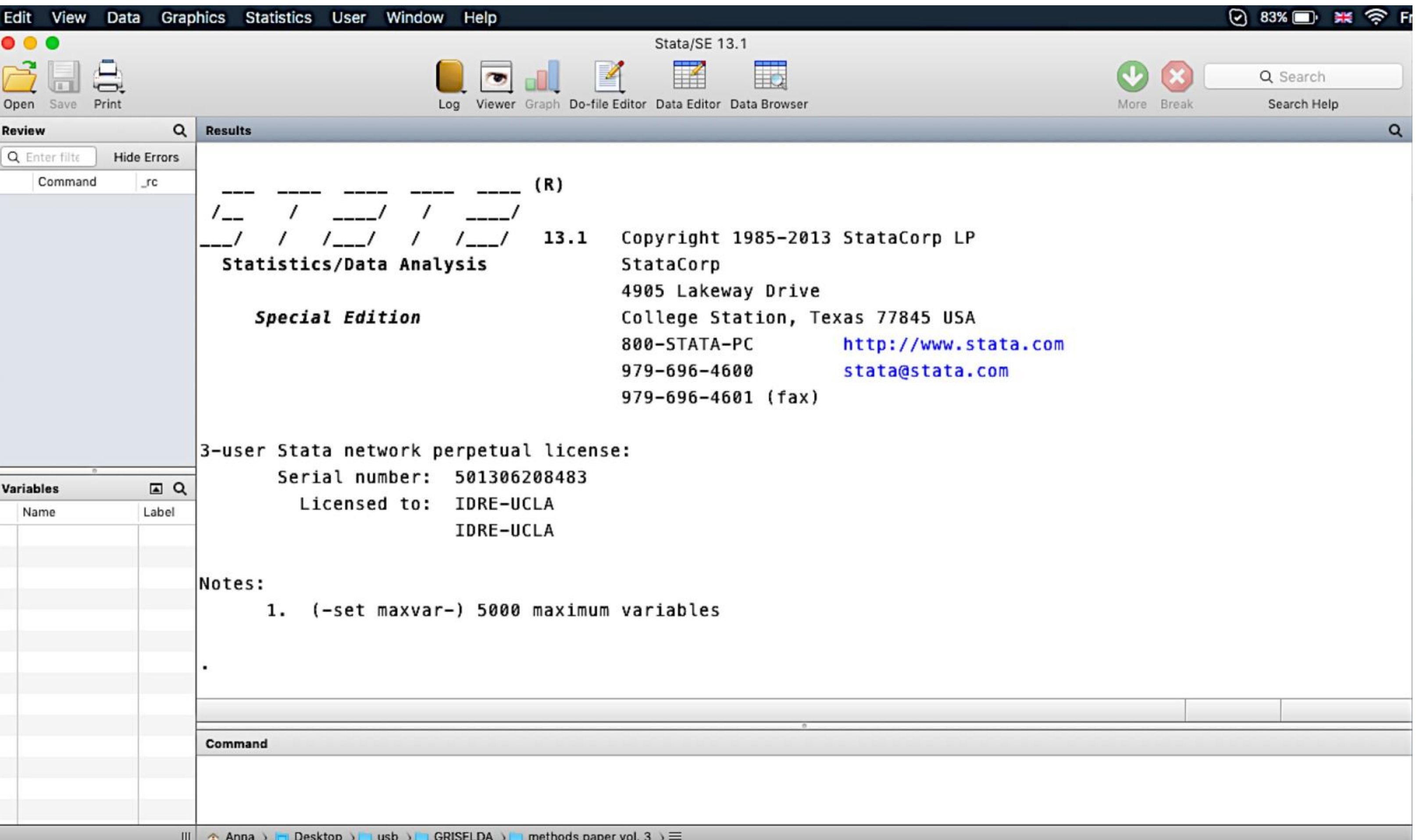


Pairwise meta-analysis – binary data



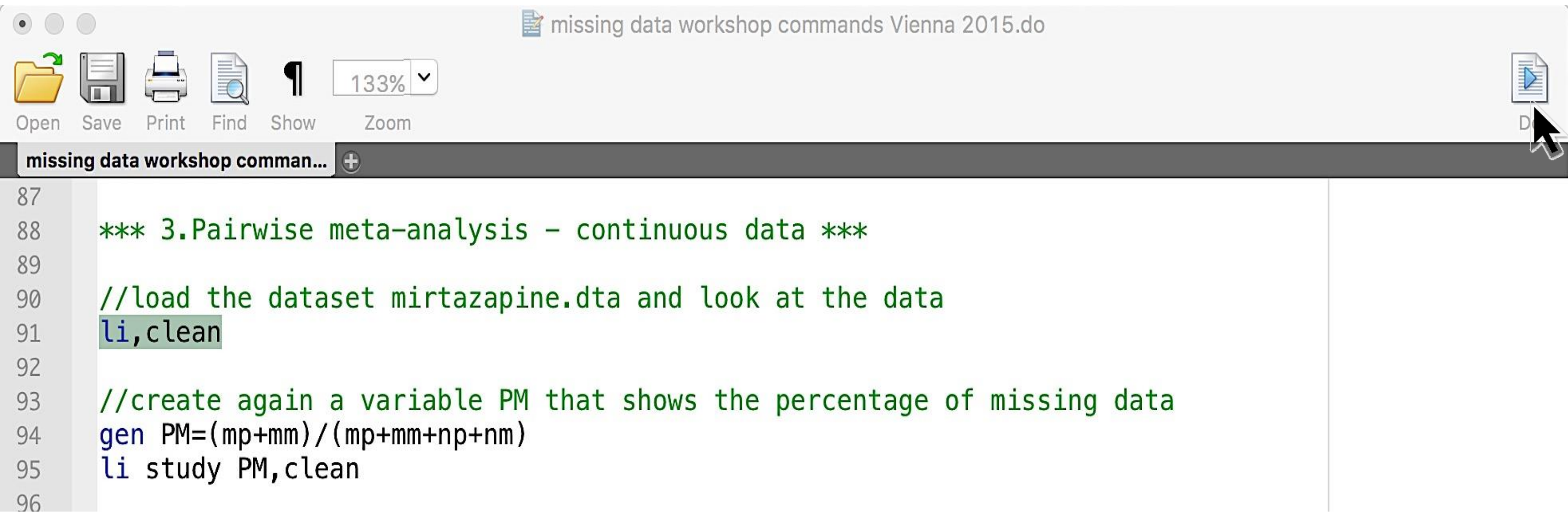
Pairwise meta-analysis – continuous data

- ✓ Load the dataset "mirtazapine.dta"



Pairwise meta-analysis – continuous data

- ✓ Look at the dataset

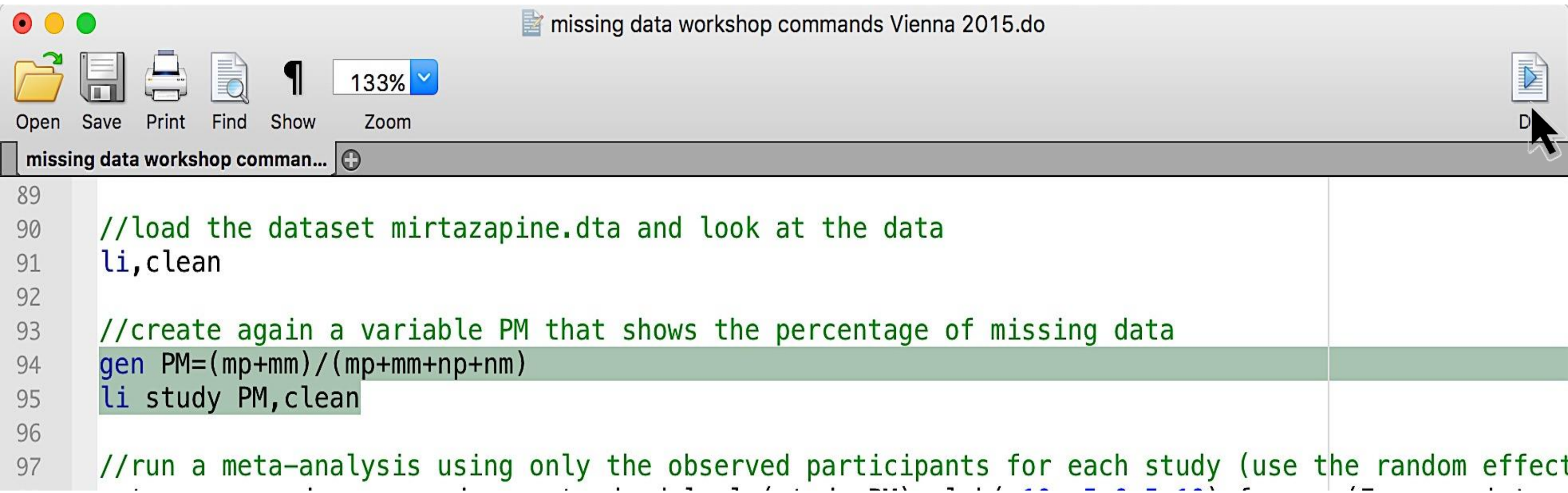


Pairwise meta-analysis – continuous data

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

Pairwise meta-analysis – continuous data

✓ Find the percentage of missing data for each study



Pairwise meta-analysis – continuous data

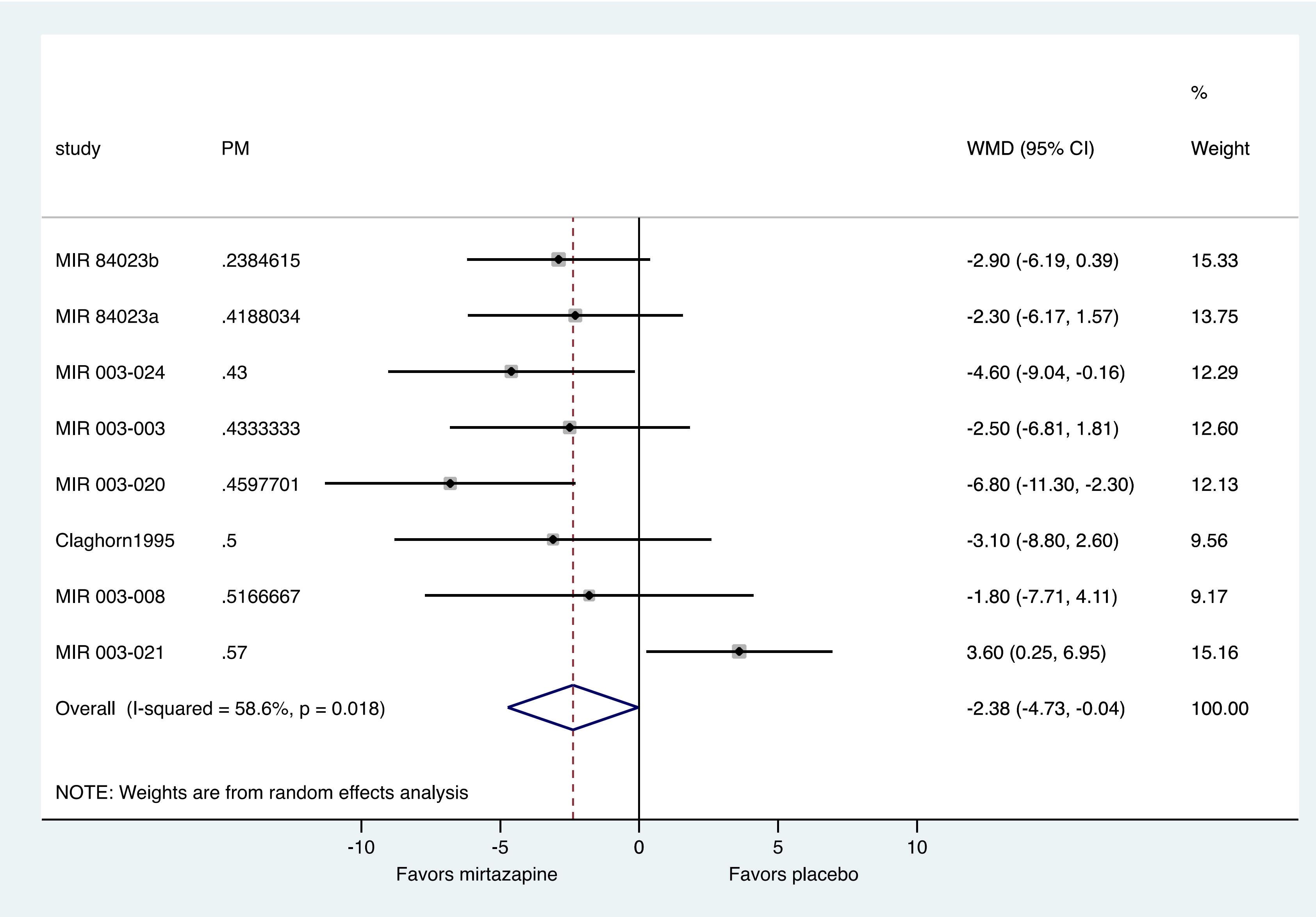
study	PM
Claghorn1995	.5
MIR 003-003	.4333333
MIR 003-008	.5166667
MIR 003-020	.4597701
MIR 003-021	.57
MIR 003-024	.43
MIR 84023a	.4188034
MIR 84023b	.2384615

Pairwise meta-analysis – continuous data

✓ Run a meta-analysis using only the observed participants for each study (use the random effects model); that is the available-case analysis (ACA)

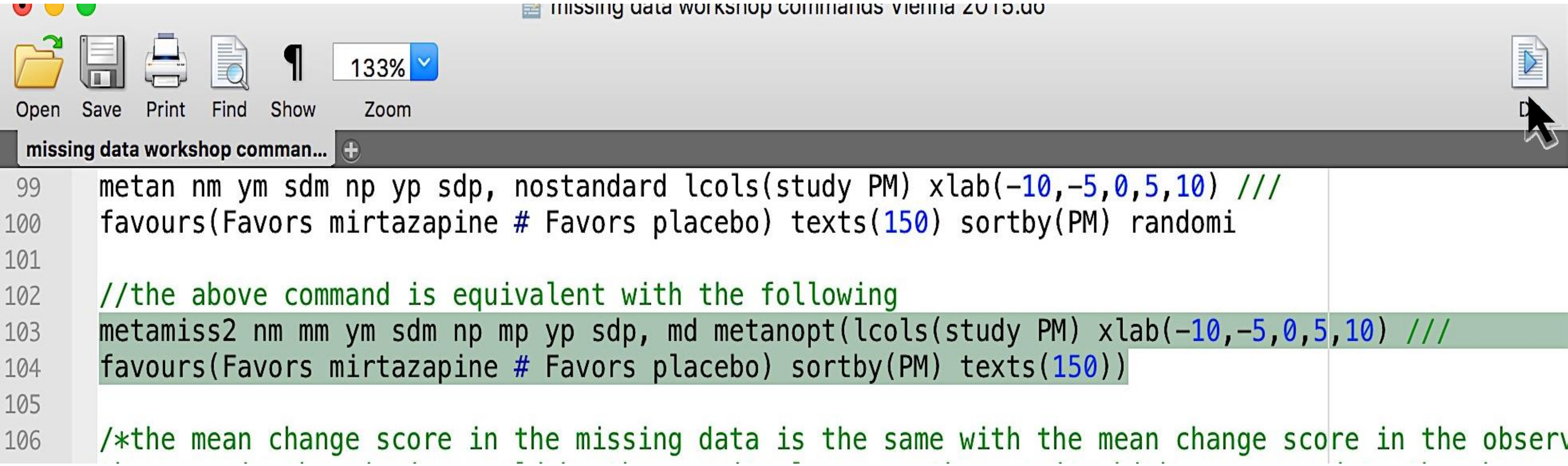
```
/*run a meta-analysis using only the observed participants for each study
(use the random effects model); that is the available-case analysis (ACA)*/
metan nm ym sdm np yp sdp, nostandard lcols(study PM) xlab(-10,-5,0,5,10) ///
favours(Favors mirtazapine # Favors placebo) texts(150) sortby(PM) randomi

//the above command is equivalent with the following
```

Pairwise meta-analysis – continuous data

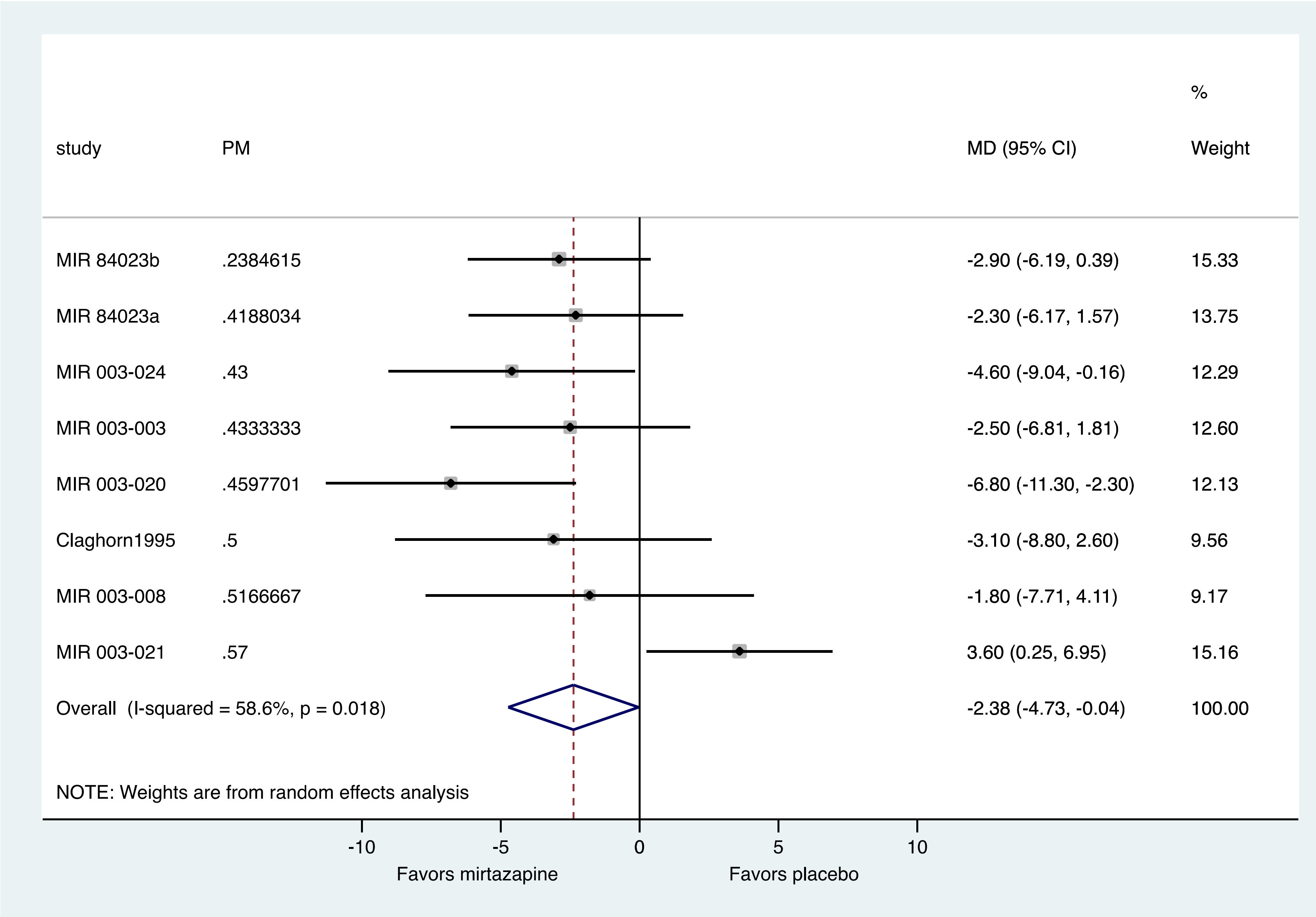


Pairwise meta-analysis – continuous data

✓ Run the same analysis with `metamiss2` instead of `metan`



Pairwise meta-analysis – continuous data



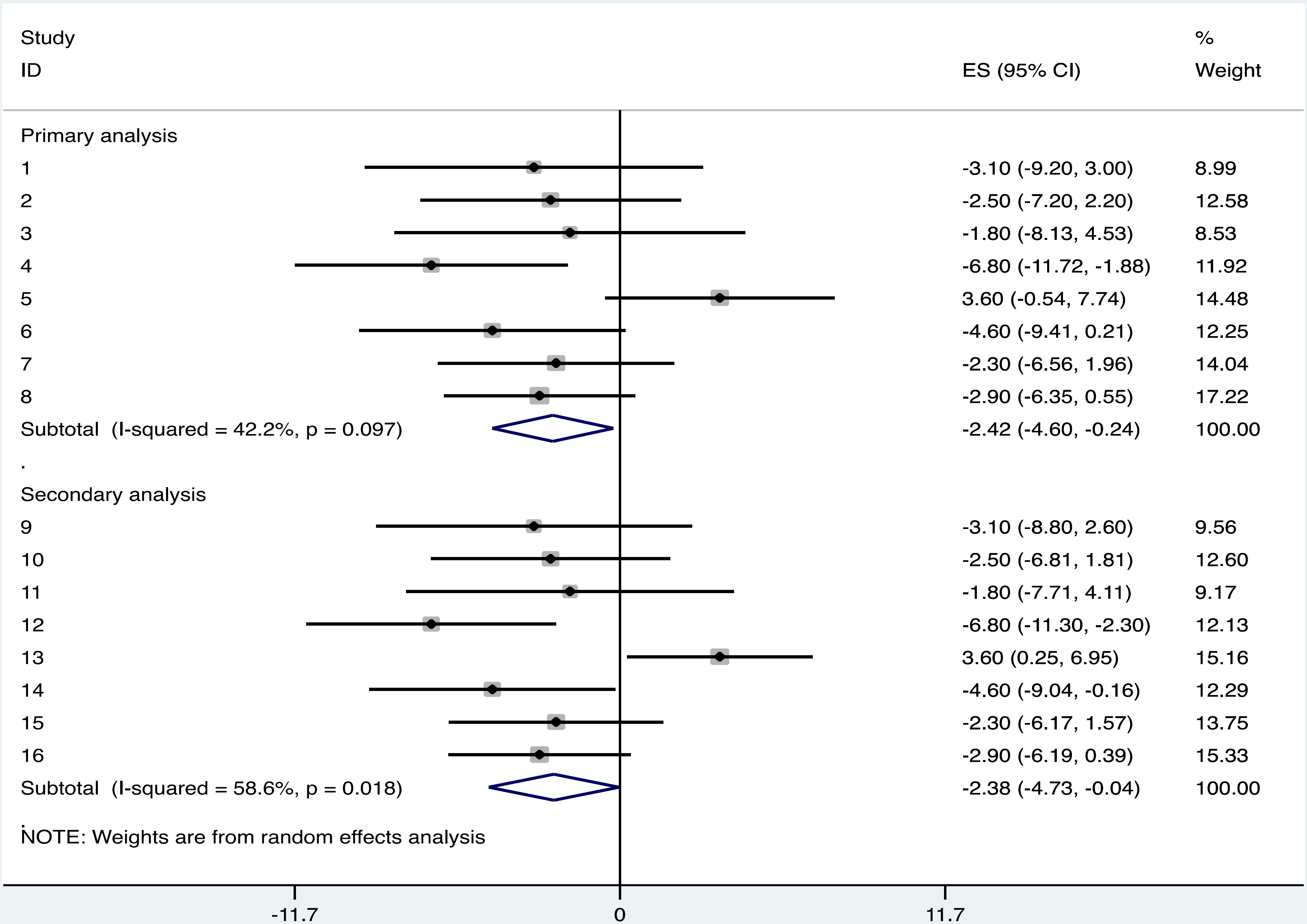
Pairwise meta-analysis – continuous data

✓ The mean change score in the missing data is the same with the mean change score in the observed data for both arms, but with some uncertainty; the mean in the missing could be three units lower or three units higher compared to the observed, compare the results with ACA

missing data workshop comman...

```
/*the mean change score in the missing data is the same with the mean change score in the
observed data for both arms, but with some uncertainty; the mean in the missing could be three
units lower or three units higher compared to the observed, compare the results with ACA*/
di 3/1.96
metamiss2 nm mm ym sdm np mp yp sdp, impmean(0) impsd(1.53) compare(impmean(0) impsd(0)) ///
md metanopt(lcols(study PM) sortby(PM) xlab(-10,-5,0,5,10) ///
favours(Favors mirtazapine # Favors placebo))

//run a sensitivity analysis showing how the uncertainty in IMDOM impacts the summary effect
```

Pairwise meta-analysis – continuous data

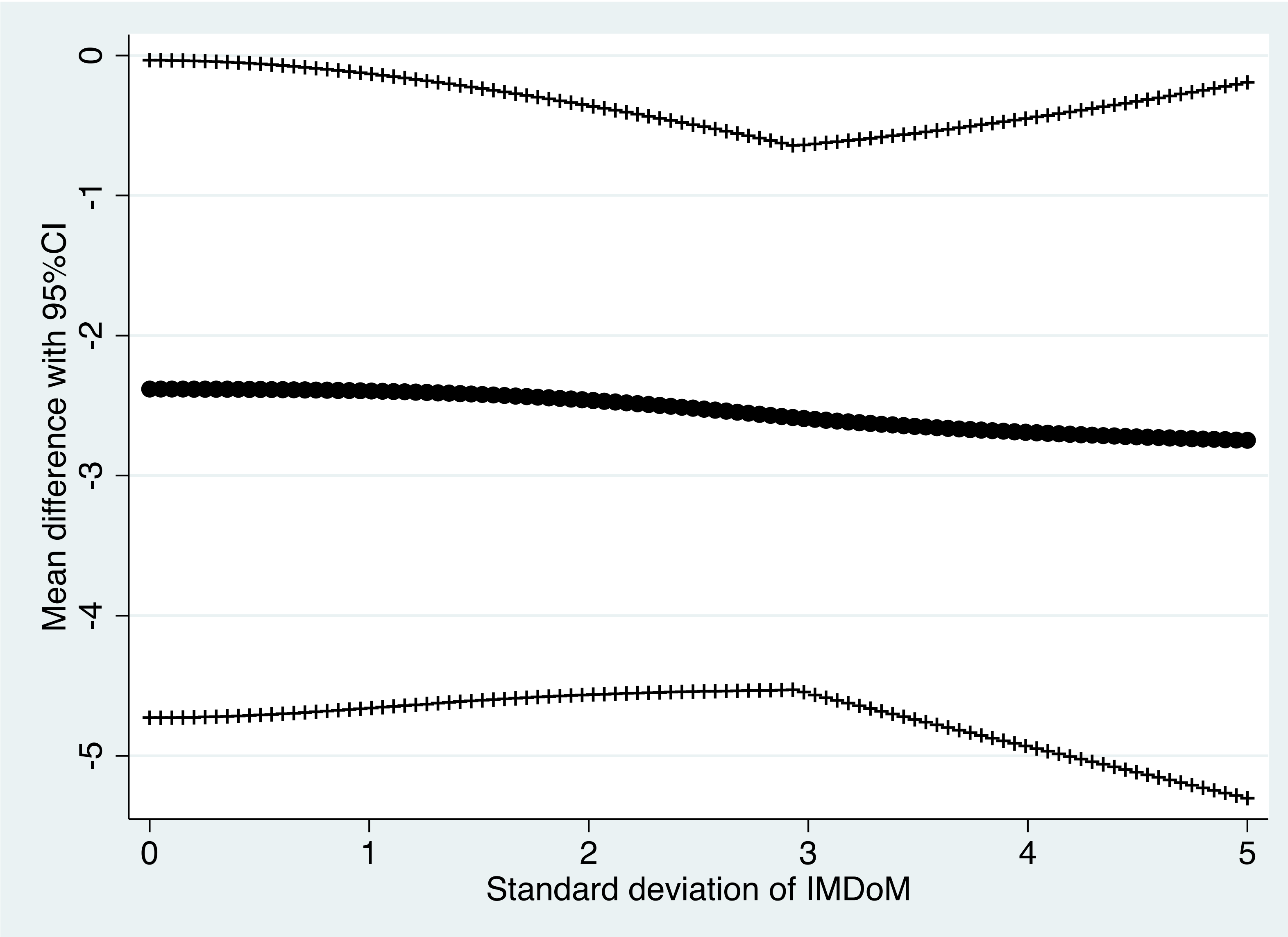


Pairwise meta-analysis – continuous data

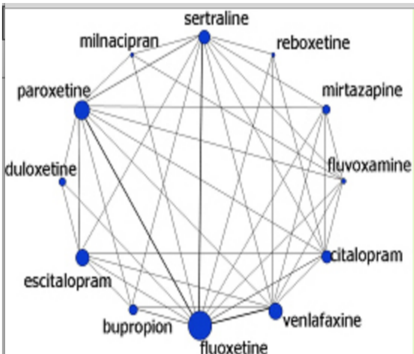
✓ Run a sensitivity analysis showing how the uncertainty in IMDOM impacts the summary effect

```
metamiss2 nm mm ym sdm np mp yp sdp, impmean(0) impsd(1.53) compare(impmean(0) impsd(0)) ///
md metanopt(lcols(study PM) sortby(PM) xlab(-10,-5,0,5,10) ///
favours(Favors mirtazapine # Favors placebo))

//run a sensitivity analysis showing how the uncertainty in IMDOM impacts the summary effect
metamiss2 nm mm ym sdm np mp yp sdp, md sensitivity
```

Pairwise meta-analysis – continuous data



Updates on www.mtm.uoi.gr



Multiple-Treatments Meta-Analysis

A Framework for Evaluating and Ranking Multiple Healthcare Technologies

CLOSE INFO

- HOME
- TUTORIAL and BIBLIOGRAPHY
- HOW TO DO AN MTM
- IMMA ERC starting Grant
- RESEARCH and PUBLICATIONS
- Stata routines for Network Meta-Analysis
- Material from Publications (software and protocols)
- Meta-analysis methods

Using Stata for Standard Pairwise and Network Meta-Analysis

We hope you will find the material presented useful. If you use the Stata routines please kindly cite the articles that describe them:

For the metamiss2 command

Mavridis D, White IR, Higgins JPT, Cipriani A, Salanti G. Allowing for uncertainty due to missing continuous outcome data in pairwise and network meta-analysis. Stat Med 2015; 34: 721–741

White IR, Higgins JPT, Wood A. Allowing for uncertainty due to missing data in meta-analysis—Part 1: Two-stage methods. Stat Med 2008; 27:711–727

For the network graphs package

Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical Tools for Network Meta-Analysis in STATA. PLoS One. 2013 Oct 3;8(10):e76654.

Chaimani A, Salanti G. Visualizing assumptions and results in network meta-analysis: the network graphs package. Stata Journal 2015; 15(4): 905-950.

Latest News

Network meta-analysis in Stata
UPDATED to fix a bug of the netweight command. Type: . net install network_graphs, from(http://www.mtm.uoi.gr) replace to install the latest versi...

3rd Course on Network Meta-Analysis, 4-6 July 2016, Oxford
3rd Course on Network Meta-Analysis, 4-6 July 2016, Oxford How to appraise, interpret and publish a network meta-analysis A 3-day course for clinici...

A comprehensive collection of methodological articles
Download here

Plenary talk at the 20th Cochrane Colloquium