Reference based multiple imputation; for sensitivity analysis of clinical trials with missing data

Suzie Cro

MRC Clinical Trials Unit at UCL
The London School of Hygiene and Tropical Medicine

Outline

- Reference based multiple imputation; asthma trial
- The mimix command
- Sensitivity analysis example 1; asthma trial
- Sensitivity analysis example 2; peer review study

Example - asthma trial

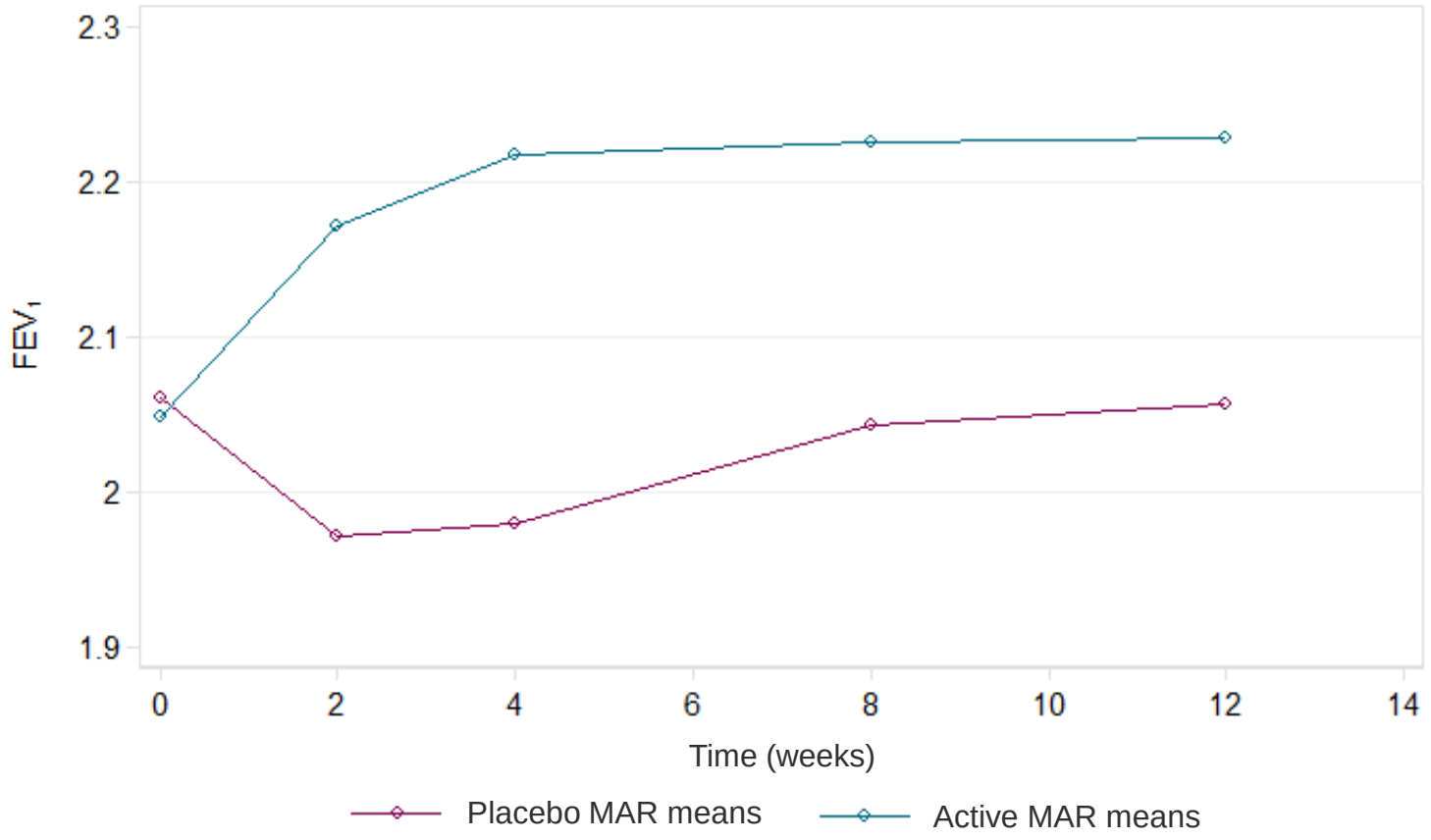
- Placebo vs. Budesonide for patients with chronic asthma
- Forced Expiratory Volume in 1 second (FEV_1) recorded at baseline, 2, 4, 8 and 12 weeks
- Primary outcome: mean treatment group difference at 12 weeks adjusted for baseline
- Only 38/90 Placebo and 72/90 Budesonide were observed at 12 weeks

Busse et al. (1998)

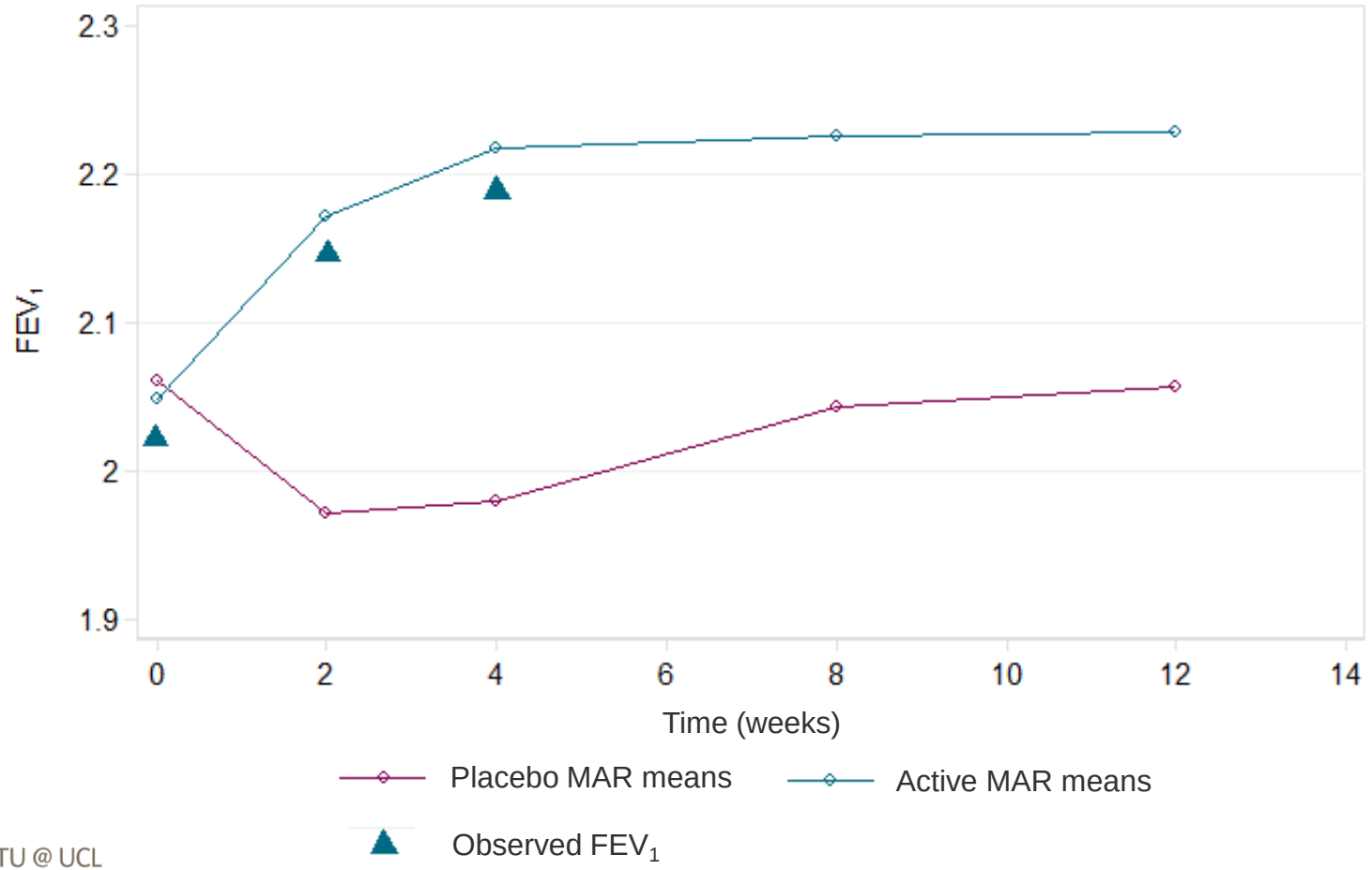
Example - asthma trial

- Any analysis must make an *untestable* assumption about the unobserved data
- Wrong assumption → biased treatment estimate
- Primary analysis – Missing-at-Random (MAR)
- A set of analyses where the missing data is handled in different ways as compared to the primary analysis should be undertaken

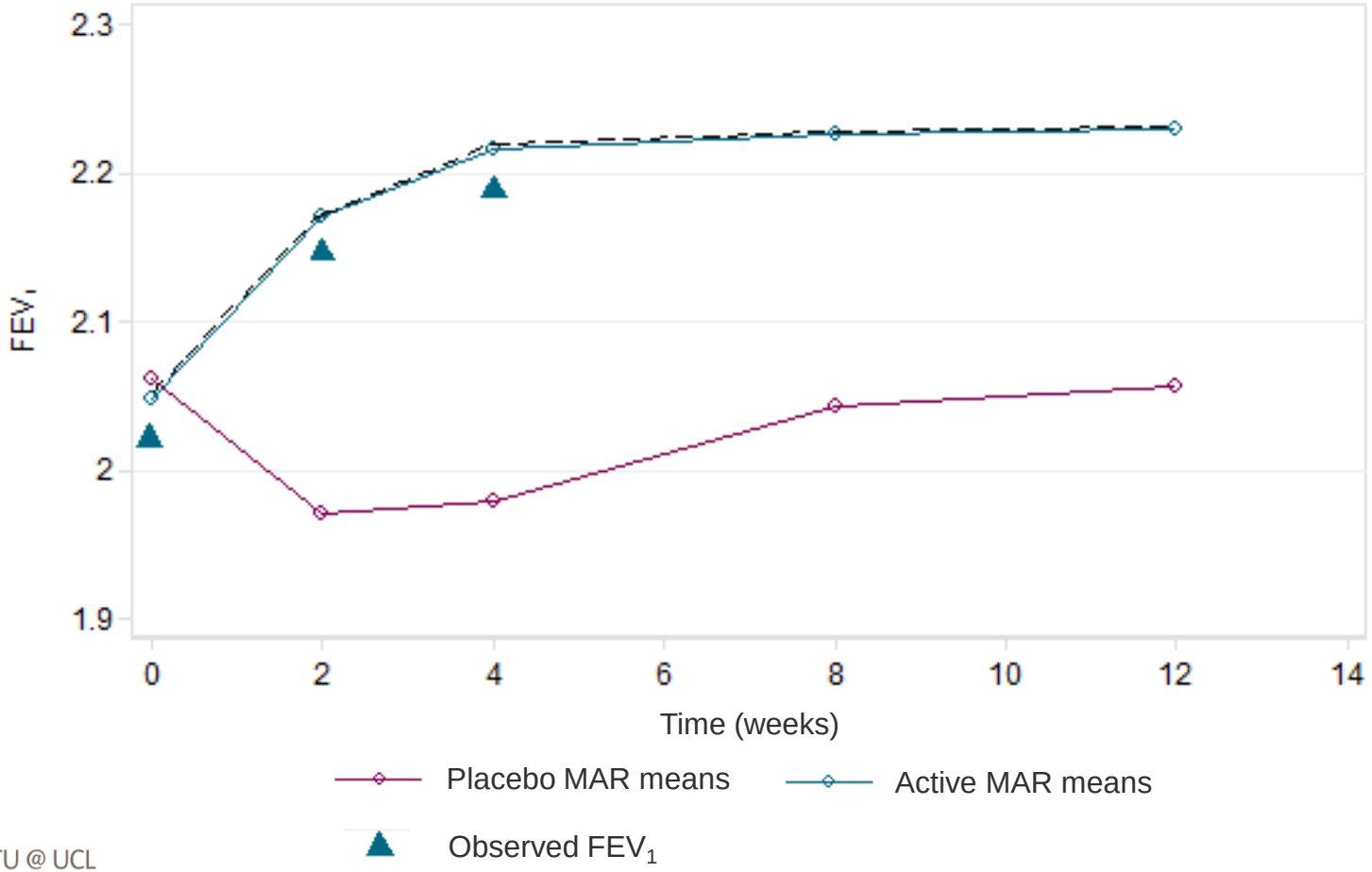
Example - asthma trial - MAR



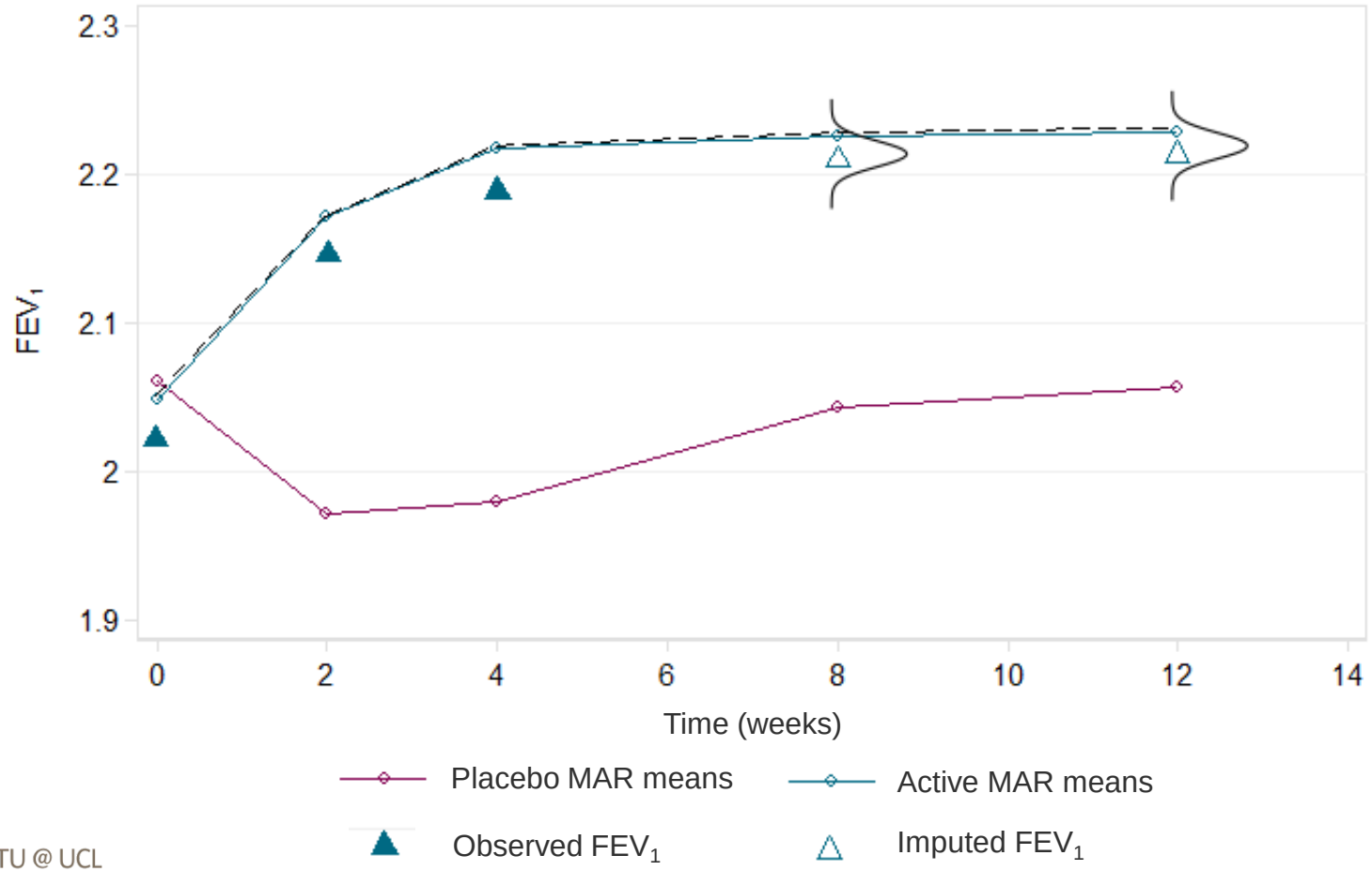
Example - asthma trial - MAR



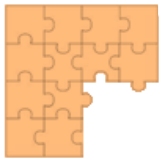
Example - asthma trial - MAR



Example - asthma trial - MAR

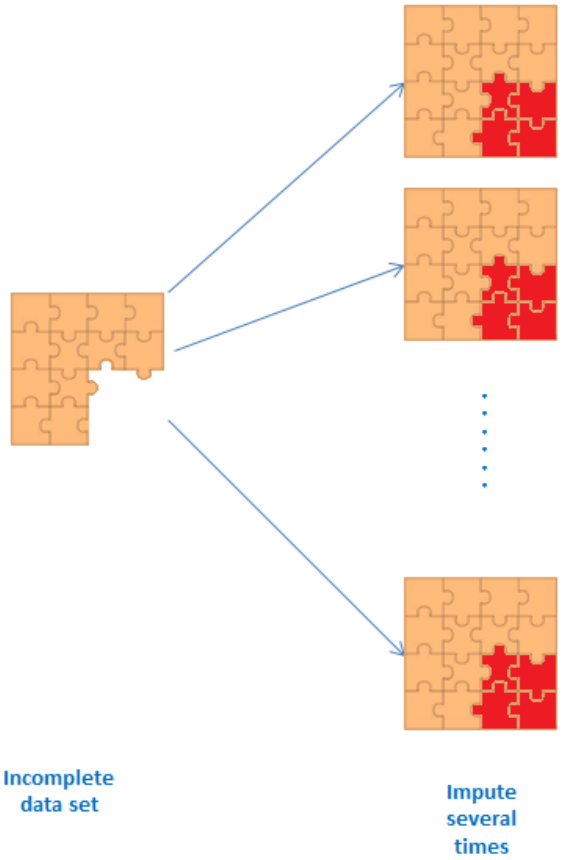


Multiple imputation

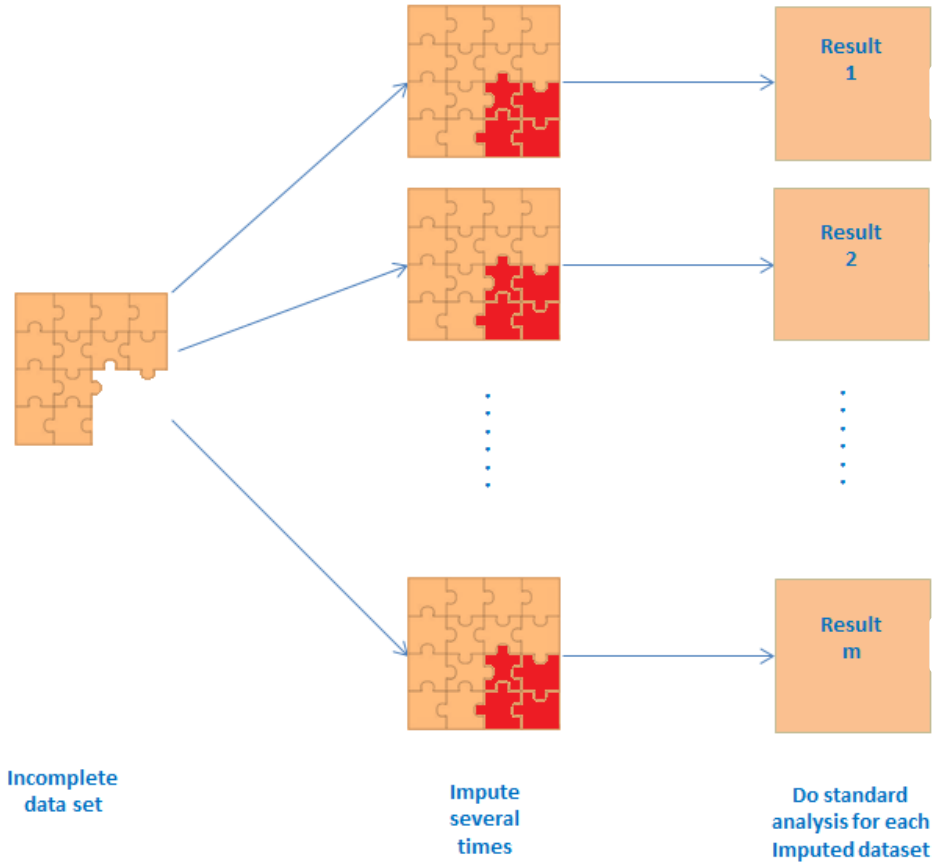


Incomplete
data set

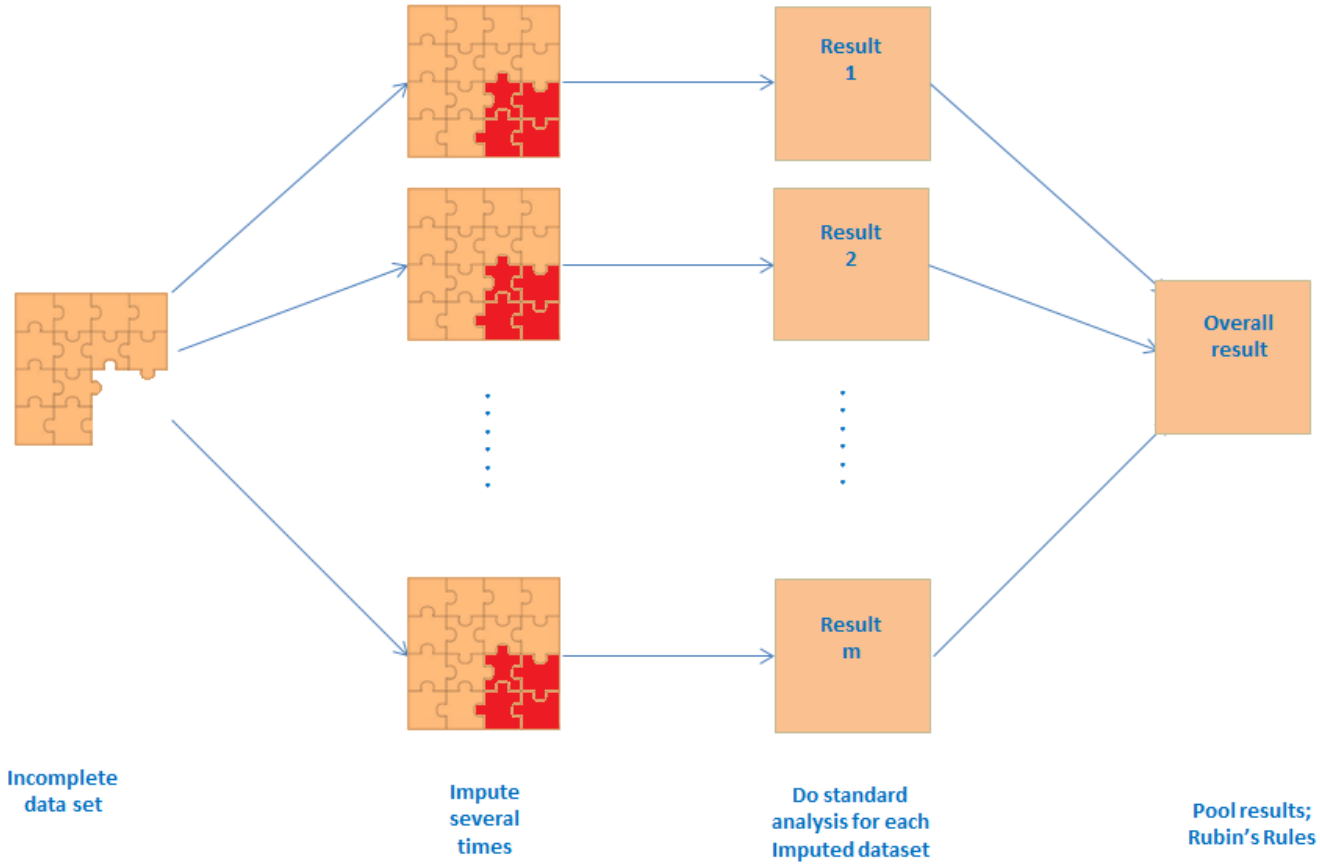
Multiple imputation



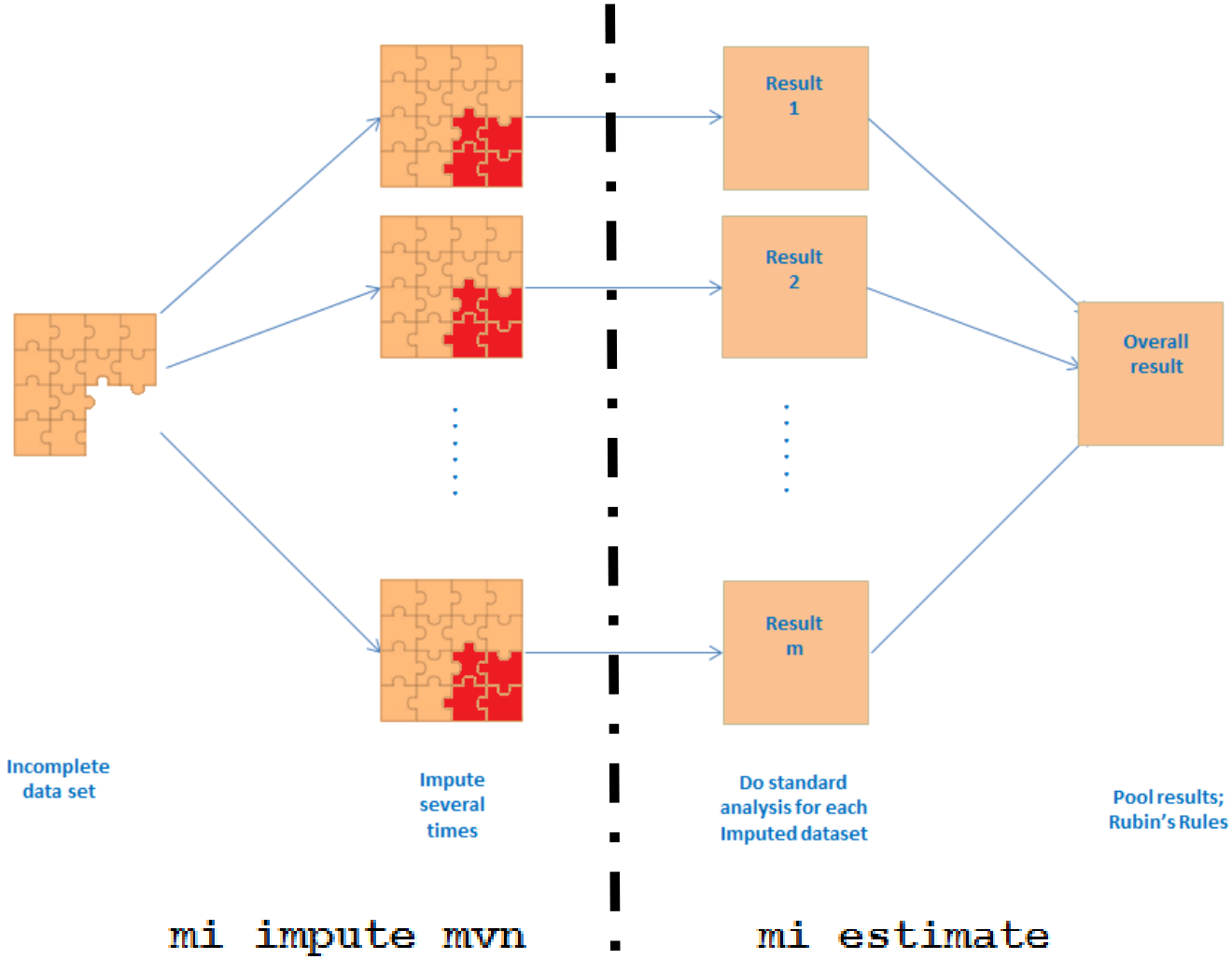
Multiple imputation



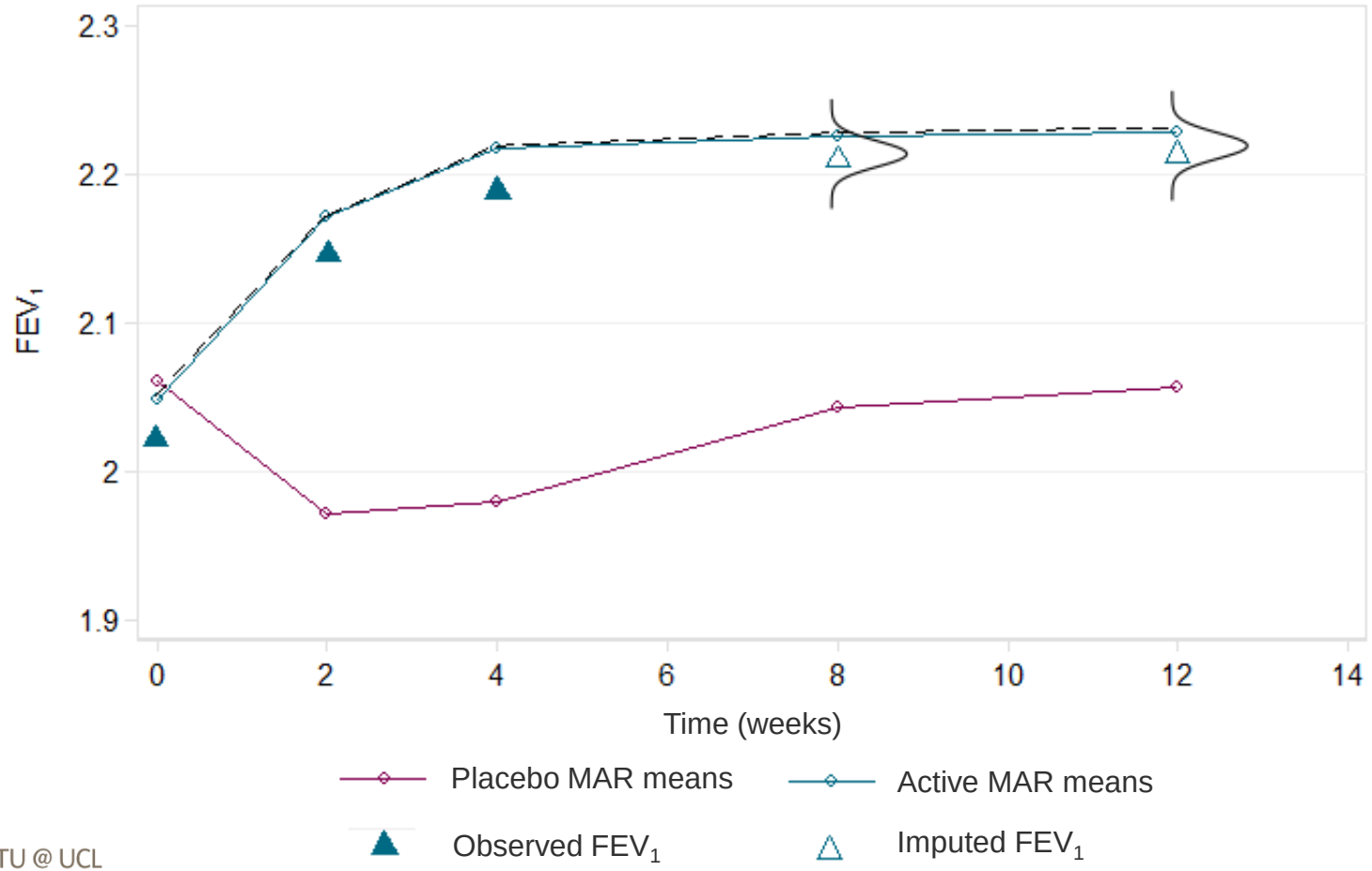
Multiple imputation



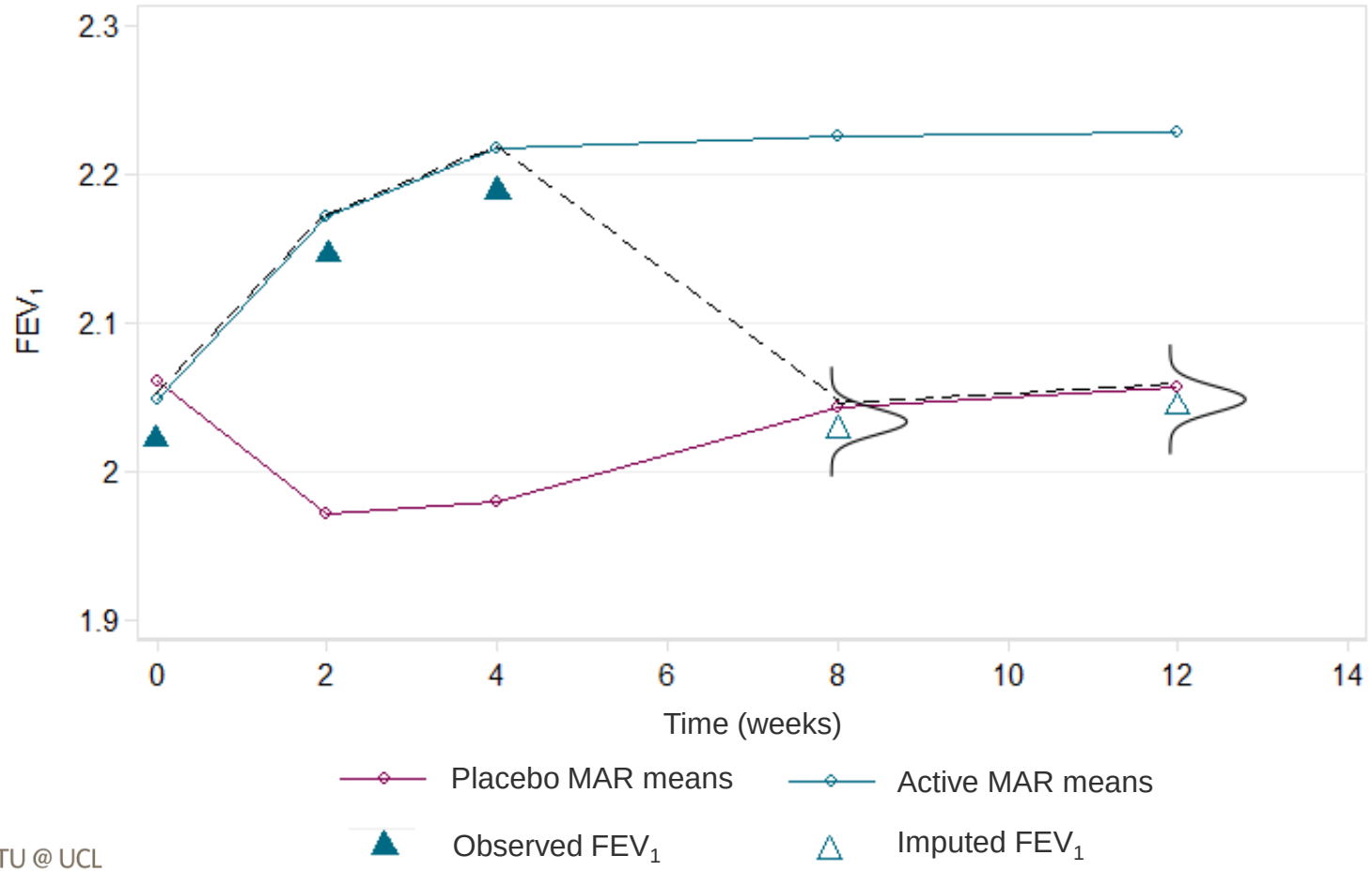
Multiple imputation



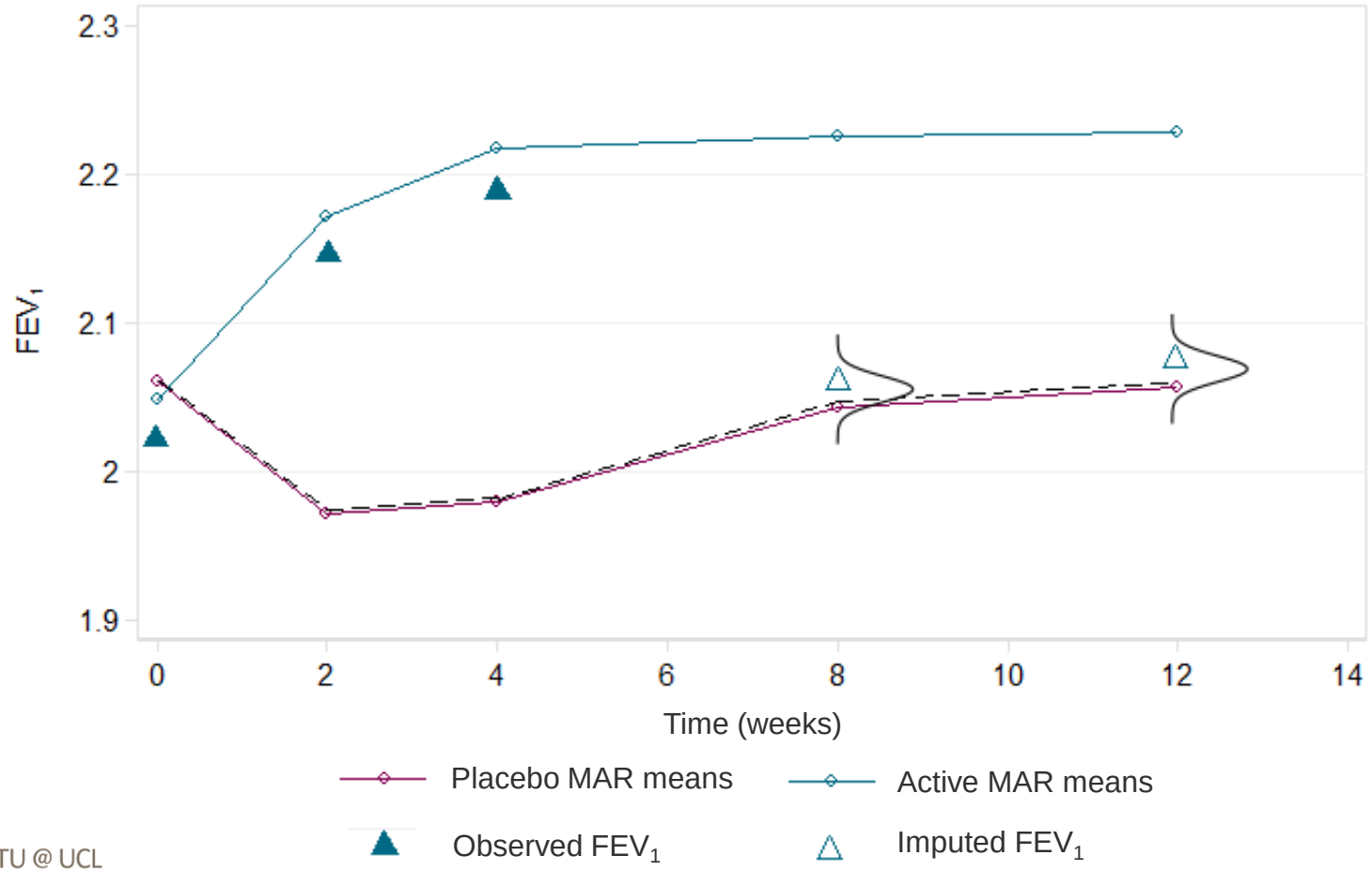
Asthma trial - MAR



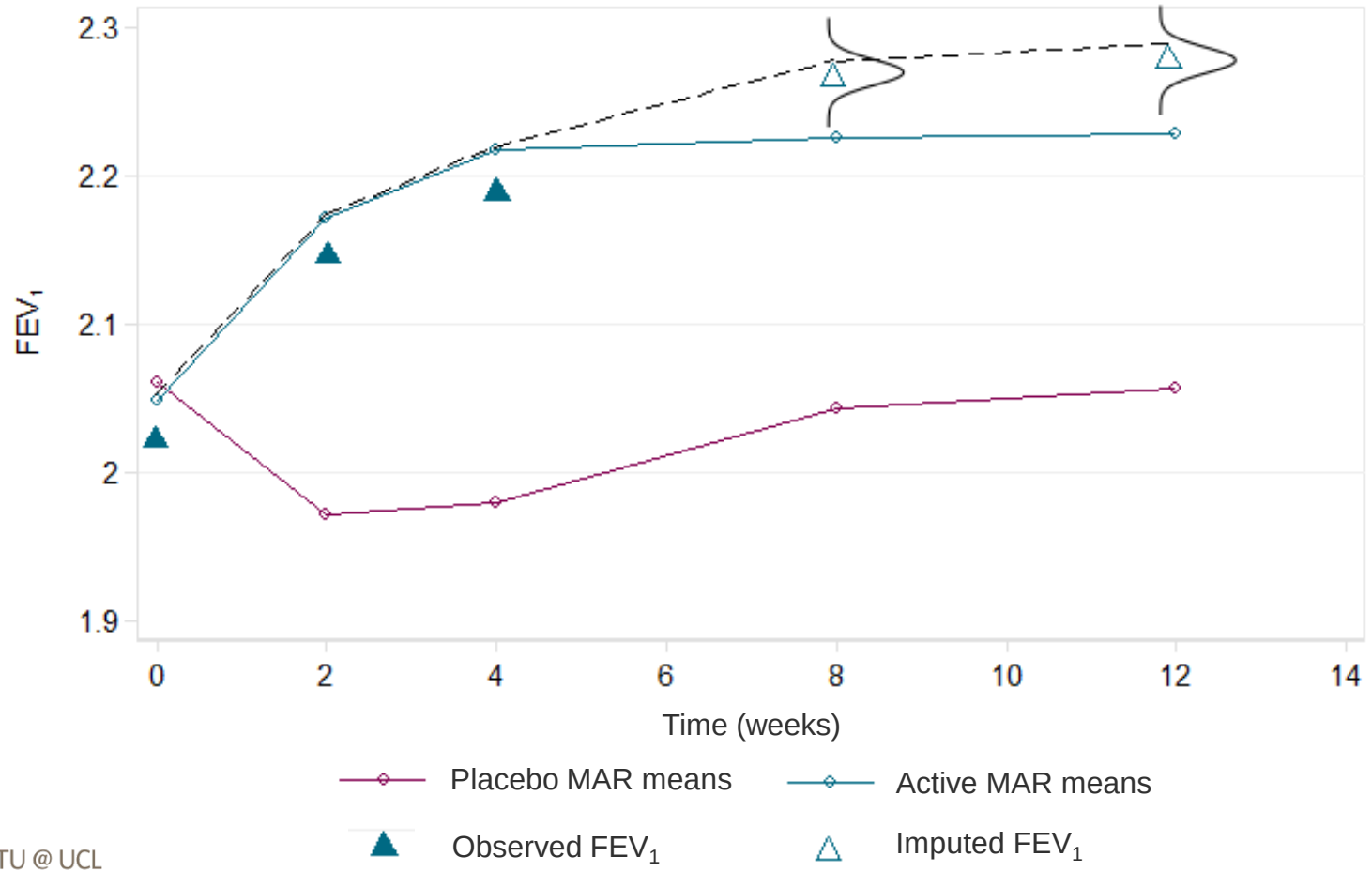
Asthma trial - Jump to reference



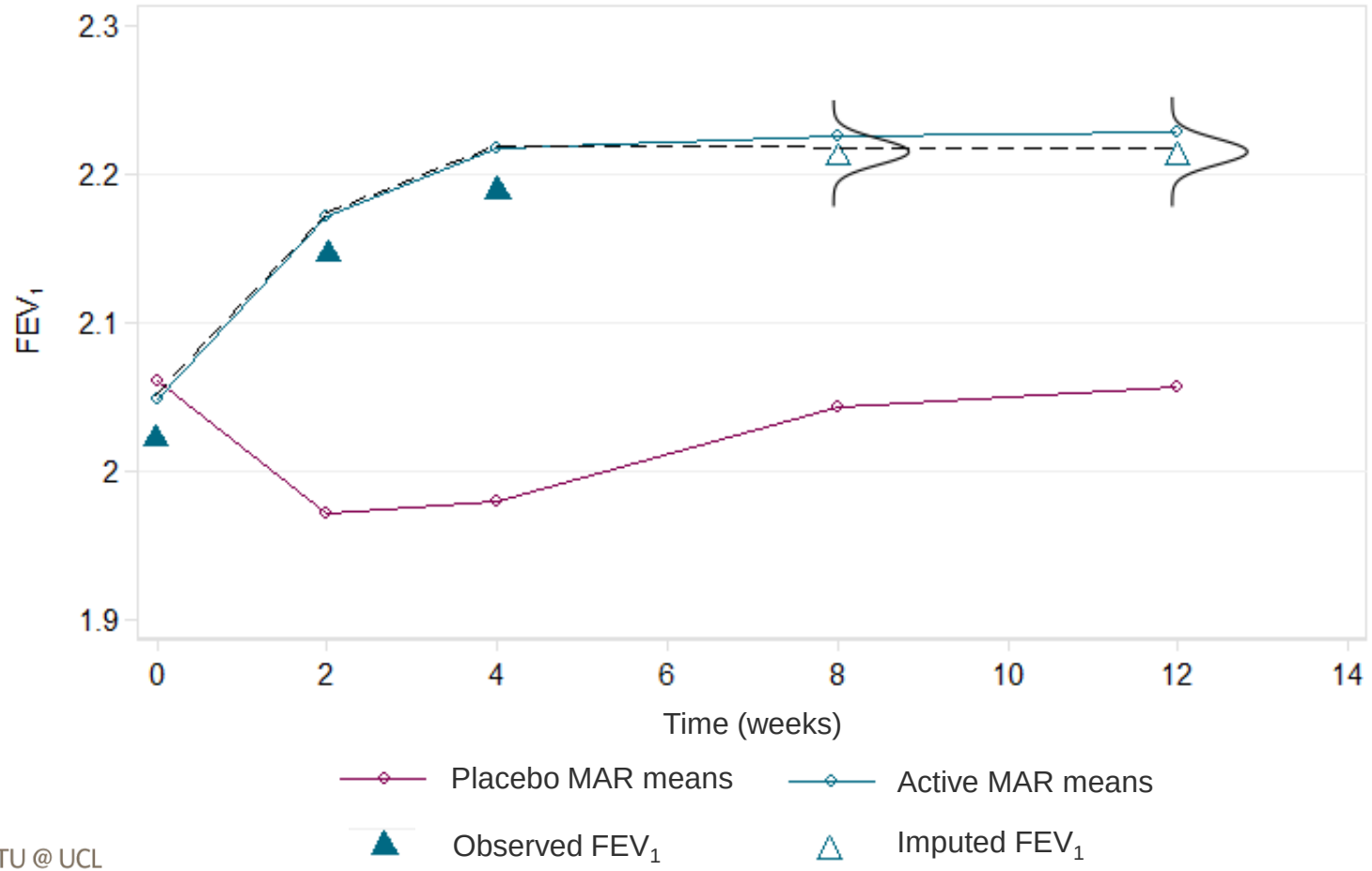
Asthma trial - Copy reference



Asthma trial - Copy increments in reference



Asthma trial - Last mean carried forward



Reference based sensitivity analysis

- Comparison of results under different reference based assumptions allows us to determine the robustness of results

- Interim missing observations may often be imputed under on-treatment MAR, or under one of the outlined assumptions

mimix

- The mimix command conducts multiple imputation under the 5 reference based assumptions
- Optionally allows users to conduct analysis with two in-built analysis options; regress or mixed
- Syntax:

```
mimix depvar treatvar, id(varname) time(varname) [ clear covariates(varlist)
interim(string) iref(string) method(string) methodvar(varname) mixed
refgroup(string) refgroupvar(varname) regress saving(filename[,replace])
burnbetween(#) burnin(#) m(#) seed(#) ]
```

Asthma trial

```
. describe

Contains data from asthma.dta
  obs:           732
 vars:             5                          30 Aug 2016 13:14
 size:        11,712
-------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------------------------
id              int     %8.0g                 Patient ID
fev             float   %9.0g                 FEV1 (L)
time            byte    %9.0g                 Measurement time (weeks)
base            double  %12.0g                Baseline FEV1 (L)
treat           byte    %8.0g      treat1     Randomised treatment assignment
-------------------------------------------------------------------------------
Sorted by: id

. label list
treat1:
           2 Placebo
           3 Active

. list in 37/40, noobs sepby(id)
```

id	fev	time	base	treat
5030	.85	2	1.14	Placebo
5030	1.51	4	1.14	Placebo
5030	.	8	1.14	Placebo
5030	.	12	1.14	Placebo

Asthma trial

```
. mimix fev treat, id(id) time(time) method(mar) covariates(base) regress m(50)
> clear seed(101)
```

Asthma trial

```
. mimix fev treat, id(id) time(time) method(mar) covariates(base) regress m(50)
> clear seed(101)
Performing imputation procedure for group 1 of 2...
Performing imputation procedure for group 2 of 2...
Performing regress procedure ...

i.treat           _Itreat_2-3         (naturally coded; _Itreat_2 omitted)

Multiple-imputation estimates                 Imputations       =         50
Linear regression                             Number of obs     =        183
                                              Average RVI       =     0.4106
                                              Largest FMI       =     0.3495
                                              Complete DF       =        180
DF adjustment:   Small sample                 DF:     min       =      91.39
                                                      avg       =      99.15
                                                      max       =     105.79
Model F test:       Equal FMI                 F(   2,   149.8)  =      40.69
Within VCE type:          OLS                 Prob > F          =     0.0000
```

fev	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itreat_3	.3230728	.1042794	3.10	0.002	.1163241	.5298215
base	.7240691	.0861441	8.41	0.000	.5531672	.8949709
_cons	.3959986	.1971734	2.01	0.048	.0043602	.787637

```
Imputed dataset now loaded in memory
Imputed data created in variable fev using mar
```

Asthma trial

```
. mimix fev treat, id(id) time(time) method(j2r) refgroup(2) interim(mar) ///
> covariates(base) regress m(50) clear seed(101)
```

Asthma trial

```
. mimix fev treat, id(id) time(time) method(j2r) refgroup(2) interim(mar) ///
> covariates(base) regress m(50) clear seed(101)
Performing imputation procedure for group 1 of 2...
Performing imputation procedure for group 2 of 2...
Performing regress procedure ...

i.treat           _Itreat_2-3         (naturally coded; _Itreat_2 omitted)

Multiple-imputation estimates                 Imputations       =         50
Linear regression                             Number of obs     =        183
                                              Average RVI       =     0.4483
                                              Largest FMI       =     0.3510
                                              Complete DF       =        180
DF adjustment:   Small sample                 DF:     min       =      91.07
                                                      avg       =     109.09
                                                      max       =     140.18
Model F test:       Equal FMI                 F(   2,   156.9)  =      32.45
Within VCE type:          OLS                 Prob > F          =     0.0000
```

fev	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Itreat_3	.2261827	.1028346	2.20	0.029	.0228754	.42949
base	.6894261	.0933944	7.38	0.000	.5040403	.8748119
_cons	.4669997	.2112431	2.21	0.030	.0473954	.8866041

```
Imputed dataset now loaded in memory
Imputed data created in variable fev using j2r
Interim missing data imputed using mar
```

Specifying the imputation method - 1

Method	Name to specify in method()
Missing at random (MAR)	mar
Jump to reference	j2r
Last mean carried forward	lmcf
Copy increments in reference	cir or ciir
Copy reference	cr

- For j2r, cir or cr also require refgroup() to specify the reference group

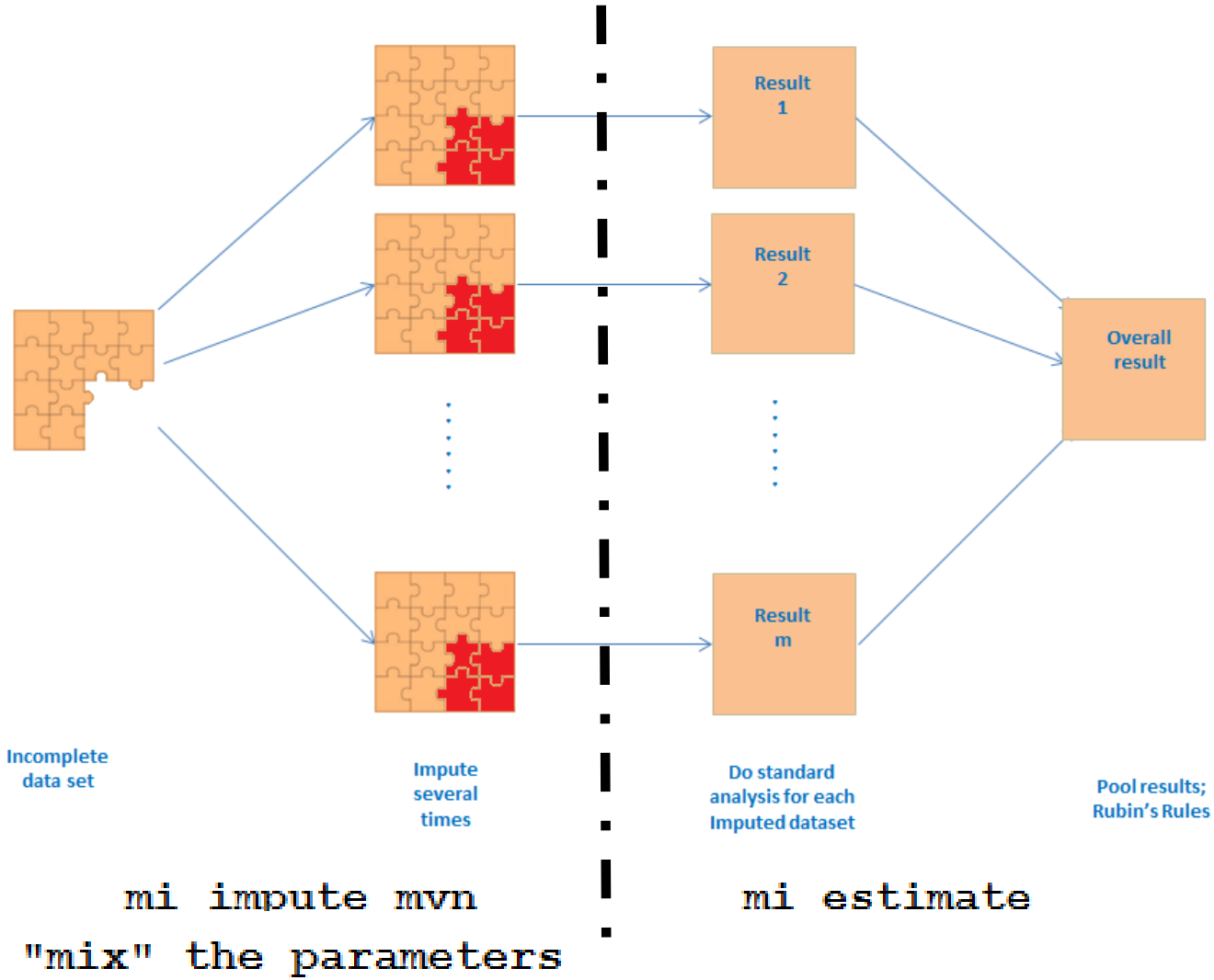
Asthma trial - results

Analysis	Treat Est.	Std. Err.	P-value
Primary – MAR	0.323	0.104	0.002
Last mean carried forward	0.296	0.096	0.003
Copy placebo	0.289	0.101	0.005
Copy active	0.251	0.082	0.003
Jump to placebo	0.226	0.103	0.029
Jump to active	0.128	0.095	0.181
Copy increments in placebo	0.281	0.103	0.007
Copy increments in active	0.277	0.082	0.001

Asthma trial - results

Analysis	Treat Est.	Std. Err.	P-value
Primary – MAR	0.323	0.104	0.002
Last mean carried forward	0.296	0.096	0.003
Copy placebo	0.289	0.101	0.005
Copy active	0.251	0.082	0.003
Jump to placebo	0.226	0.103	0.029
Jump to active	**0.128**	**0.095**	**0.181**
Copy increments in placebo	0.281	0.103	0.007
Copy increments in active	0.277	0.082	0.001

mimix - behind the scenes…



Example 2 - Reviewer study

- Schroter et al. (2004) performed a single blind RCT among BMJ reviewers to compare:
  - no training
  - self-taught training
- Participants sent a baseline paper to review (paper 1)
- 2-3 months later sent second paper to review
- Quality of review measured by the mean (2 raters) Review Quality Instrument, range 1 to 5

Example 2 - Reviewer study

- Quality of baseline review:

	No intervention			Self training		
	n	mean	SD	n	mean	SD
Returned paper 2	162	2.65	0.81	120	2.80	0.62
Did not return paper 2	11	3.02	0.50	46	2.55	0.75

Example 2 - Reviewer study

- Quality of baseline review:

	No intervention			Self training		
	n	mean	SD	n	mean	SD
Returned paper 2	162	2.65	0.81	120	2.80	0.62
Did not return paper 2	11	3.02	0.50	46	2.55	0.75

- Primary analysis – MAR assumption

- What if participants who did not return paper 2 behaved like the no intervention group?

Example 2 - Reviewer study

```
. describe

Contains data from reviewer.dta
  obs:           339
 vars:             5                          30 Aug 2016 13:50
 size:         6,780
-------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------------------------
id              float   %9.0g                 Reviewer identifier
inter           float   %19.0g     inter1     Training package
base            float   %9.0g                 Paper 1 (baseline) mean review quality
resp            float   %9.0g                 Paper 2 (response) mean review quality
time            float   %9.0g
-------------------------------------------------------------------------------
Sorted by:

. label list
inter1:
           0 no-training
           1 self-taught

. list in 6, noobs sepby(id)
```

id	inter	base	resp	time
25	no-training	1.714286	2.928571	2

Example 2 - Reviewer study

```
. mimix resp inter, id(id) time(time) covariates(base) method(mar) m(50) clear seed(23) regress
Performing imputation procedure for group 1 of 2...
Performing imputation procedure for group 2 of 2...

Performing regress procedure ...

i.inter            _Iinter_0-1         (naturally coded; _Iinter_0 omitted)

Multiple-imputation estimates                     Imputations       =         50
Linear regression                                 Number of obs     =        339
                                                  Average RVI       =     0.2180
                                                  Largest FMI       =     0.1840
                                                  Complete DF       =        336
DF adjustment:   Small sample                     DF:     min       =     230.76
                                                          avg       =     235.32
                                                          max       =     243.00
Model F test:       Equal FMI                     F(   2,   297.5)  =      35.58
Within VCE type:          OLS                     Prob > F          =     0.0000
```

resp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iinter_1	.2393743	.0702048	3.41	0.001	.1010545	.377694
base	.3610957	.0477203	7.57	0.000	.2670724	.4551189
_cons	1.60063	.1352016	11.84	0.000	1.334313	1.866947

```
Imputed dataset now loaded in memory
Imputed data created in variable resp using mar
```

Example 2 - Reviewer study

```
. mimix resp inter, id(id) time(time) covariates(base) method(cr) refgroup(0) m(50) clear seed(23)
> regress
Performing imputation procedure for group 1 of 2...
Performing imputation procedure for group 2 of 2...

Performing regress procedure ...

i.inter            _Iinter_0-1         (naturally coded; _Iinter_0 omitted)

Multiple-imputation estimates                     Imputations       =         50
Linear regression                                 Number of obs     =        339
                                                  Average RVI       =     0.1883
                                                  Largest FMI       =     0.1647
                                                  Complete DF       =        336
DF adjustment:   Small sample                     DF:     min       =     242.89
                                                          avg       =     252.09
                                                          max       =     258.48
Model F test:       Equal FMI                     F(   2,   306.6)  =      35.22
Within VCE type:          OLS                     Prob > F          =     0.0000
```

resp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_Iinter_1	.171676	.0686811	2.50	0.013	.0364214	.3069306
base	.3707551	.0471337	7.87	0.000	.2779121	.4635982
_cons	1.574795	.1331916	11.82	0.000	1.312516	1.837073

```
Imputed dataset now loaded in memory
Imputed data created in variable resp using cr
```

Example 2 - Reviewer study

Analysis	Treat Est.	Std. Err.	P-value
Primary – MAR	0.239	0.070	0.001
Copy no intervention	0.172	0.069	0.013

- The intervention effect is slightly reduced under copy no intervention but it remains statistically significant

Specifying the imputation method - 2

- For individual specific imputation methods use methodvar(*varname*) option
- Where *varname* defines the imputation method for each individual – must be constant over time
- refgroupvar(*varname*) defines individual specific reference group

Acknowledgements

- Adaptation of a SAS macro written by James Roger
- Thanks to Tim Morris for his comments and editions which helped to improve the programme
- James Carpenter, Mike Kenward

Carpenter JR, Roger JH, Kenward MG, Analysis of Longitudinal Trials with protocol deviation: a framework for relevant accessible assumptions and inference via multiple imputation, *Journal of Biopharmaceutical Statistics*, 23:1352-1371, 2013.

Cro S, Morris TP, Kenward MG, Carpenter JR, Reference-based sensitivity analysis via multiple imputation for longitudinal trials with protocol deviation, *Stata Journal*, 16:2:443-463, 2016.