

Contributions to CAMIS PHUSE project in 2025: Tobit regression, reference-based multiple imputation and tipping point analysis

Paper AS07
Phuse Connect EU 2025, Hamburg, Germany
17 November 2025

Yannick Vandendijck (J&J), Christina Sotto (J&J), Sarah Brosens (KULeuven), Christina E Fillmore (GSK), Lyn Taylor (Parexel)

Senior Principal Scientist Biostatistics
Johnson & Johnson, Belgium

The views and opinions expressed in this presentation are those of the speakers and do not necessarily reflect the views or positions of any entities they represent

Johnson & Johnson
Innovative Medicine



Me on July 7, 2023 at 14h27



QC results gave a different one-sided p-value between SAS[®] and R

Analysis set: Per Protocol	Placebo 7	Low Dose ¹ 6	Medium Dose ¹ 6	High Dose ¹ 10
Tobit Analysis of Variance				
N	7	6	6	9
Least Square Means (90% CI)	5.01 (4.41, 5.61)	5.20 (4.55, 5.85)	4.74 (4.08, 5.40)	2.99 (2.39, 3.59)
Least Square Difference (90% CI)		0.19 (-0.69, 1.08)	-0.27 (-1.16, 0.62)	-2.02 (-2.87, -1.17)
z-value		0.3615	-0.4990	-3.9150
p-value ²		0.6411	0.3089	<0.0001
Exact Wilcoxon Rank Sum Test				
p-value ²		0.1253	0.3141	0.0003
² One-sided p-value				

Exact Wilcoxon-Mann-Whitney Test

```
data:  AVAL by TRT01A (PLACEBO, HIGH DOSE)
Z = 3.1696, p-value = 0.0003497
alternative hypothesis: true mu is greater than 0
```

QC results gave a different one-sided p-value between SAS[®] and R

Analysis set: Per Protocol

Tobit Analysis of Variance

N

Least Square Means (90% CI)

Least Square Difference (90% CI)

z-value

p-value²

Exact Wilcoxon Rank Sum Test

p-value²

² One-sided p-value

Placebo	Low Dose ¹	Medium Dose ¹	High Dose ¹
7	6	6	10
7	6	6	9
5.01 (4.41, 5.61)	5.20 (4.55, 5.85)	4.74 (4.08, 5.40)	2.99 (2.39, 3.59)
	0.19 (-0.69, 1.08)	-0.27 (-1.16, 0.62)	-2.02 (-2.87, -1.17)
	0.3615	-0.4990	-3.9150
	0.6411	0.3089	<0.0001
	0.1253	0.3141	0.0003

Exact Wilcoxon-Mann-Whitney Test

data: AVAL by TRT01A (PLACEBO, LOW DOSE)

Z = -1.216, p-value = 0.891

alternative hypothesis: true mu is greater than 0

Exact Wilcoxon-Mann-Whitney Test

data: AVAL by TRT01A (PLACEBO, MEDIUM DOSE)

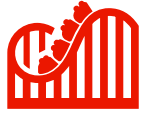
Z = -0.57143, p-value = 0.7331

alternative hypothesis: true mu is greater than 0

It would be great if there was a resource that aims to describe how analysis results in different software can be matched perfectly or knowing the source of any discrepancies



How CAMIS can help you!



A shift to R is ongoing in industry



Differences with SAS can cause uncertainty, apprehension and nervousness about the use of open-source software (e.g., R)

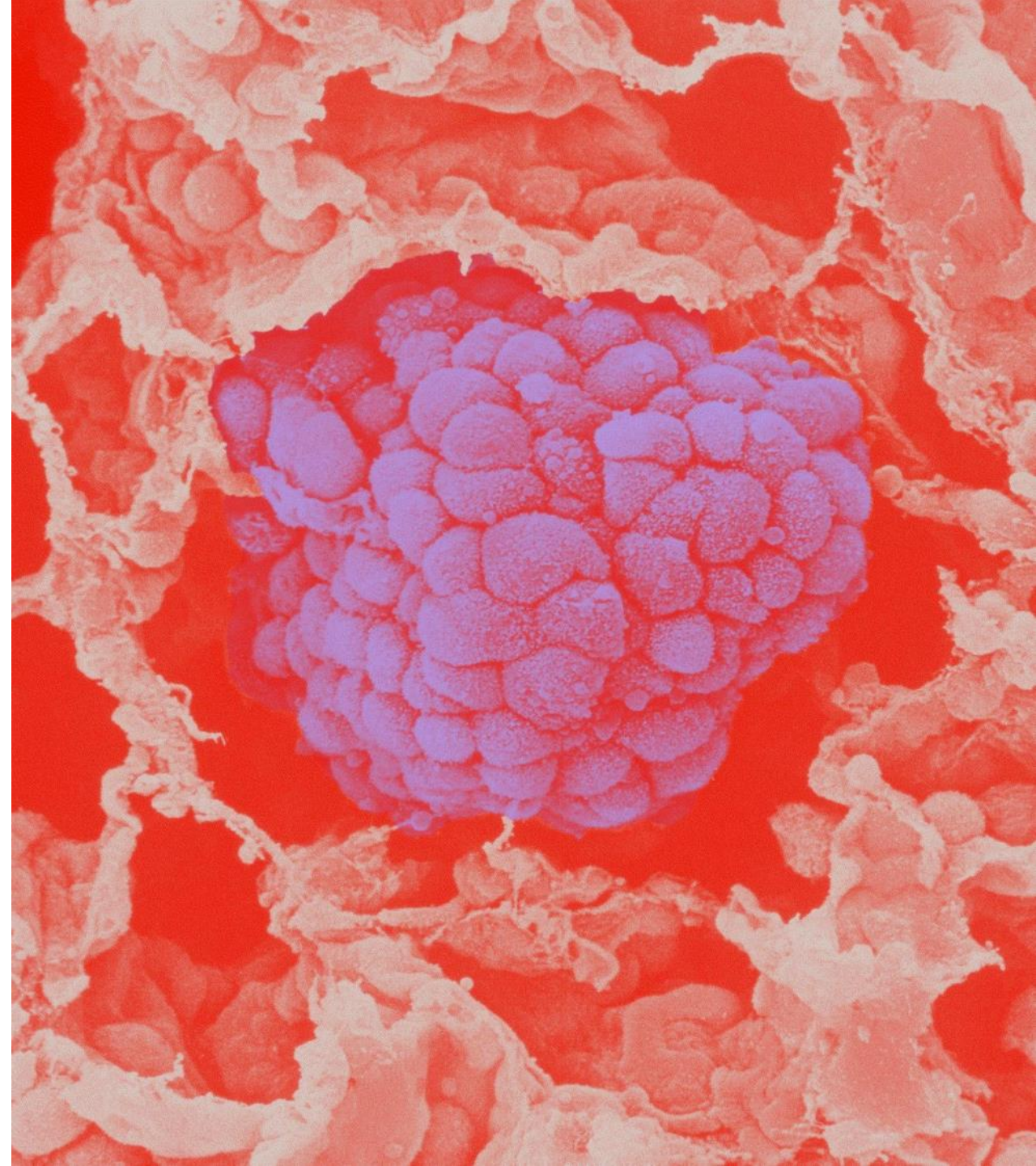


- Increase understanding and awareness of **analysis result differences** across software (R, SAS, Python)
- Understanding how to **mimic analysis results** across software is critical
- To demonstrate the methodology through **examples**, document in open-source GitHub repository

CAMIS contributes to the confidence in reliability of open-source software by understanding how analysis results can be matched perfectly or knowing the source of any discrepancies

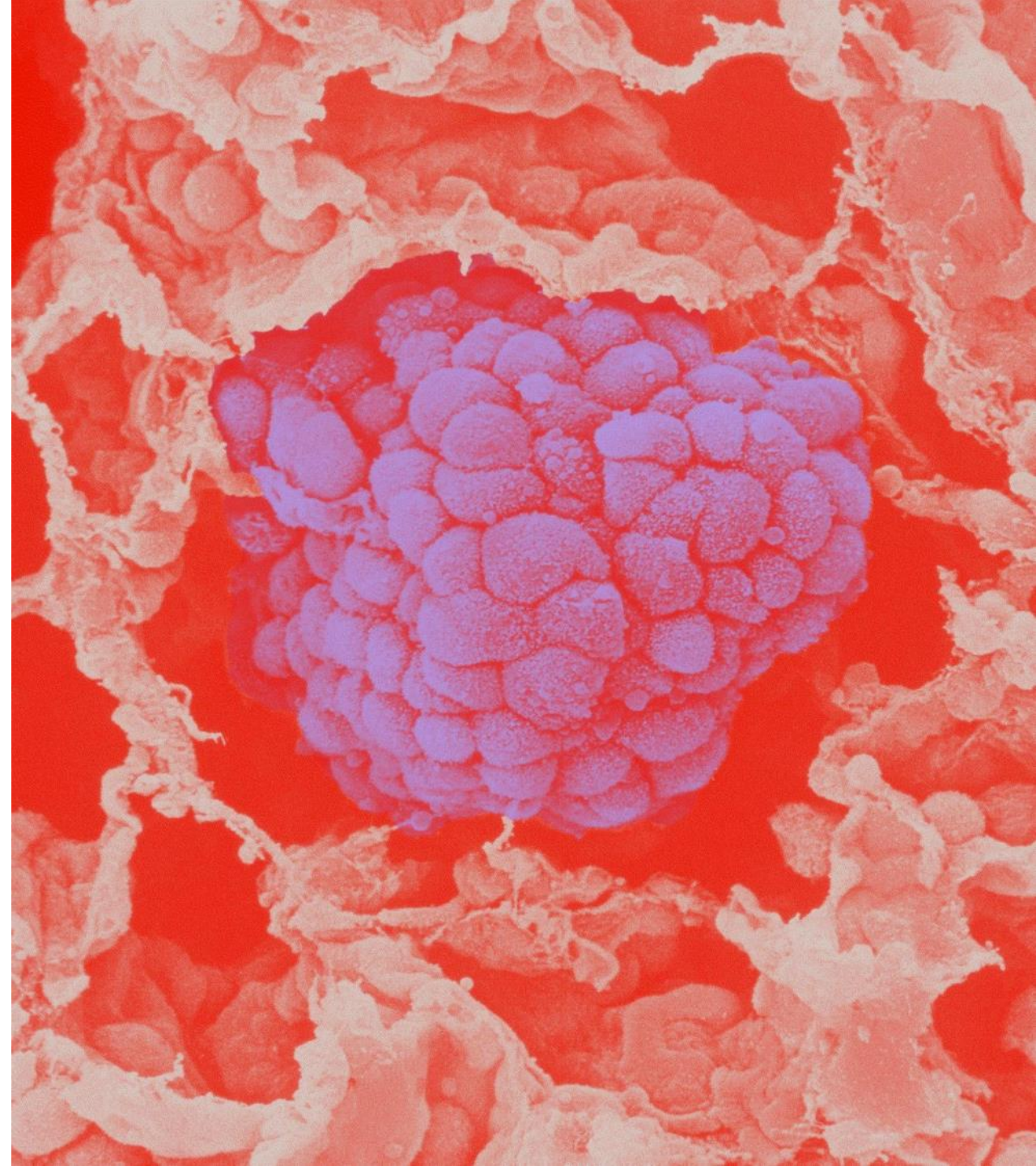
3 additions to CAMIS

1. Tobit regression
2. Reference-based multiple imputation
3. Tipping point analysis



3 additions to CAMIS

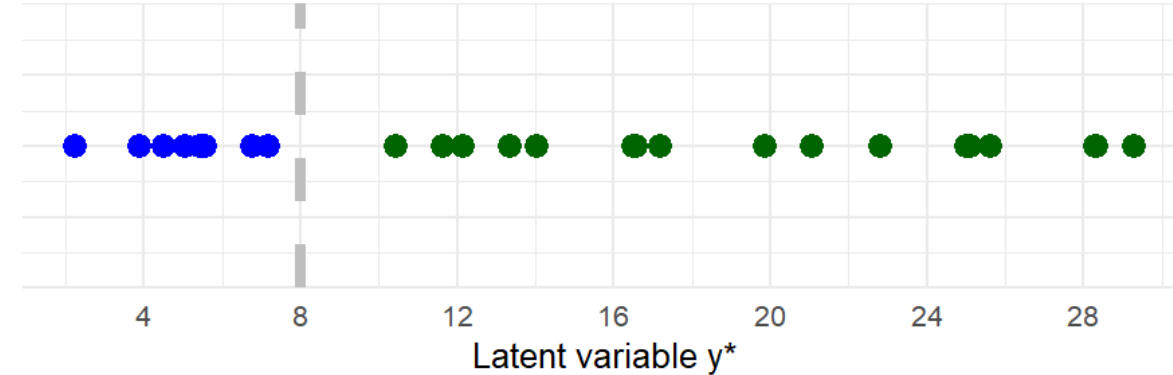
1. Tobit regression
2. Reference-based multiple imputation
3. Tipping point analysis



What is Tobit regression?

- Regression model when outcome y is censored
- Example in virology:
 - sample results below the lower limit of detection (8 copies/mL)
 - only know that sample result is <8 copies/mL, but we don't know the exact value

Example of left censoring ($\tau = 8$)

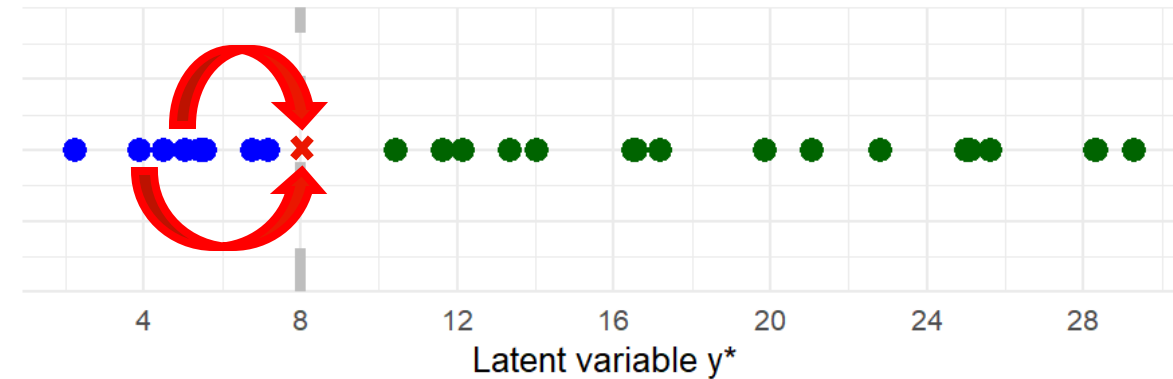


What is Tobit regression?

- Regression model when outcome y is censored
- Example in virology:
 - sample results below the lower limit of detection (8 copies/mL)
 - only know that sample result is <8 copies/mL, but we don't know the exact value
- The observed y is

$$y = \begin{cases} y^*, & \text{when } y^* \geq \tau \\ \tau, & \text{when } y^* < \tau \end{cases}$$

Example of left censoring ($\tau = 8$)



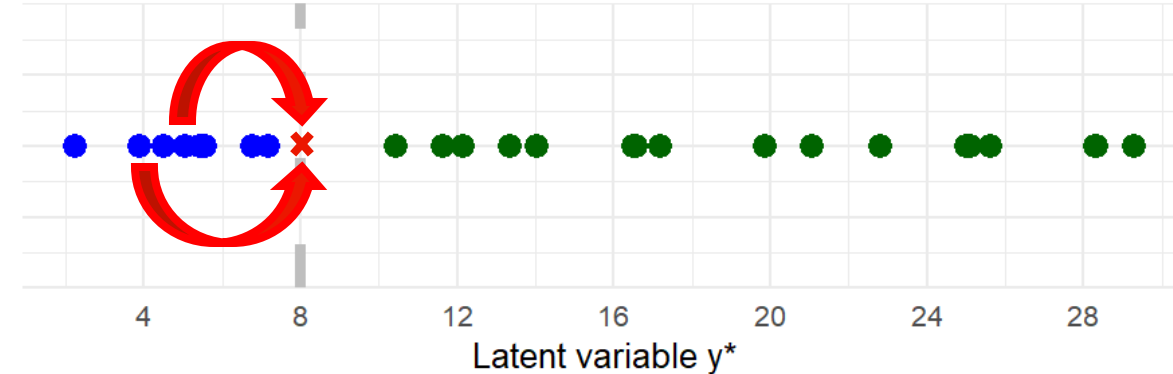
What is Tobit regression?

- Regression model when outcome y is censored
- Example in virology:
 - sample results below the lower limit of detection (8 copies/mL)
 - only know that sample result is <8 copies/mL, but we don't know the exact value

- The observed y is

$$y = \begin{cases} y^*, & \text{when } y^* \geq \tau \\ \tau, & \text{when } y^* < \tau \end{cases}$$

Example of left censoring ($\tau = 8$)



ID	ARM	Y	CENS
001	A	8.0	1
002	A	8.0	1
003	A	8.0	1
004	A	8.0	1
005	A	8.9	0
006	A	9.5	0
007	A	9.9	0

What is Tobit regression?

- Regression model when outcome y is censored
- Example in virology:
 - sample results below the lower limit of detection (8 copies/mL)
 - only know that sample result is <8 copies/mL, but we don't know the exact value

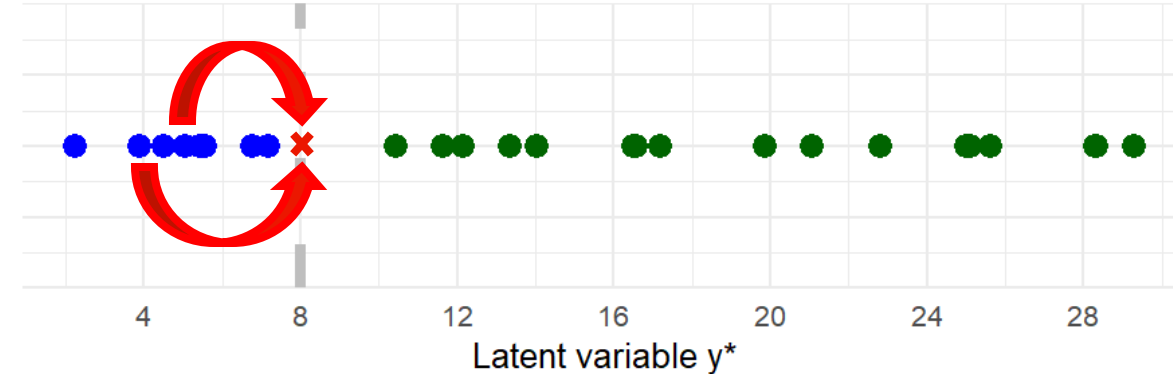
- The observed y is

$$y = \begin{cases} y^*, & \text{when } y^* \geq \tau \\ \tau, & \text{when } y^* < \tau \end{cases}$$

- Regression model on y^*

$$y_i^* = X_i\beta + \epsilon_i, \quad \text{with } \epsilon_i \sim N(0, \sigma^2)$$

Example of left censoring ($\tau = 8$)



ID	ARM	Y	CENS
001	A	8.0	1
002	A	8.0	1
003	A	8.0	1
004	A	8.0	1
005	A	8.9	0
006	A	9.5	0
007	A	9.9	0

R implementation



censReg **censReg()**

```
censReg::censReg(  
  Y ~ ARM,  
  left = 8.0,  
  data = dat_used)
```

survival **survreg()**

```
survival::survreg(  
  Surv(Y, 1-CENS,  
        type="left") ~ ARM,  
  dist = "gaussian",  
  data = dat_used)
```

VGAM **vglm()**

```
VGAM::vglm(  
  Y ~ ARM,  
  tobit(Lower = 8.0),  
  data = dat_used)
```

SAS implementation



PROC LIFEREG

```
data dat_used;
  set dat_used;
  if Y < 8.0 then lower=. ; else lower=Y;
run;

proc lifereg data=dat_used;
  class ARM;
  model (lower, Y) = ARM / d=normal;
  lsmeans ARM /cl alpha=0.05;
  estimate 'Contrast B-A' ARM 1 -1 /alpha=0.05;
run;
```

Model syntax:
if the *lower* value is missing, then the *Y* is used as a left-censored value.

Comparison

Statistic	censReg:: censReg()	survival:: survreg()	VGAM:: vglm()	PROC LIFEREG
Treatment effect	1.8225	1.8225	1.8226	
Standard error	0.8061	0.8061	0.7942	
P-value	0.0238	0.0238	0.0217	
95% CI (Wald-based)	0.2427 to 3.4024	0.2427 to 3.4024	0.2661 to 3.3791	
σ	1.7316	1.7316	1.7317	

- **VGAM::vglm** slightly different results as Iteratively Reweighted Least Squares (IRLS) algorithm used for optimization

Comparison

Statistic	censReg:: censReg()	survival:: survreg()	VGAM:: vglm()	PROC LIFEREG
Treatment effect	1.8225	1.8225	1.8226	1.8225
Standard error	0.8061	0.8061	0.7942	0.8061
P-value	0.0238	0.0238	0.0217	0.0238
95% CI (Wald-based)	0.2427 to 3.4024	0.2427 to 3.4024	0.2661 to 3.3791	0.2427 to 3.4024
σ	1.7316	1.7316	1.7317	1.7316

- **VGAM::vglm** slightly different results as Iteratively Reweighted Least Squares (IRLS) algorithm used for optimization
- **PROC LIFEREG** similar as **censReg::censReg** and **survival::survreg**
- One-sided p-values easily obtained in SAS
- In **PROC LIFEREG** and **survival::survreg** other distributional assumption are possible

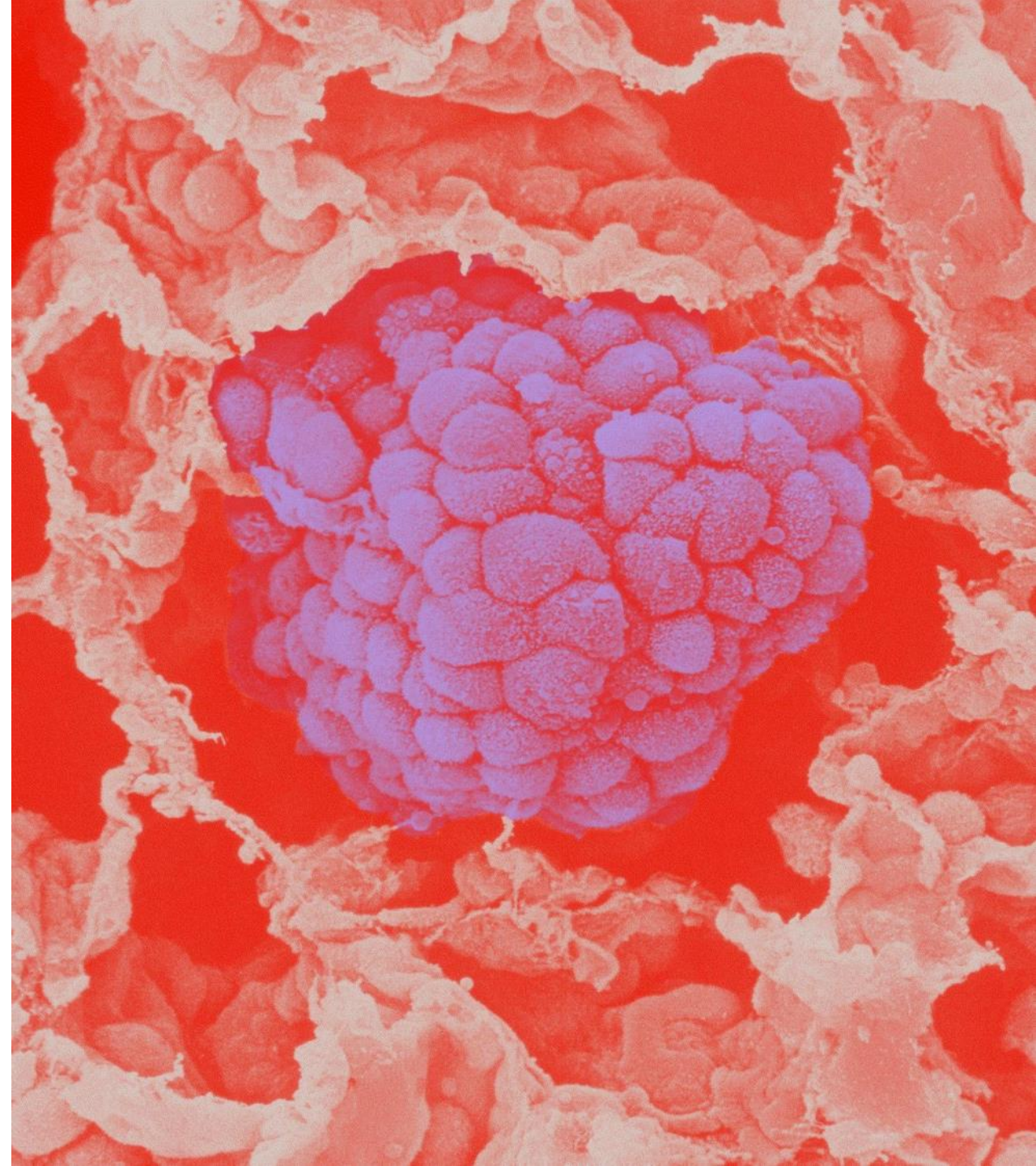
Comparison

Statistic	censReg:: censReg()	survival:: survreg()	VGAM:: vglm()
Treatment effect	1.8225	1.8225	
Standard error	0.8061	0.8061	
P-value	0.0238		
95% CI (Wald-based)	0.2427 to 3.4024		
σ			1.7316

- **VGAM::vglm()** uses **Iteratively Reweighted Least Squares (IRLS)**
- **VGAM::vglm()** and **survival::survreg**
- **VGAM::vglm()** obtained in SAS
- In **VGAM::vglm()** and **survival::survreg** other distributional assumption are possible

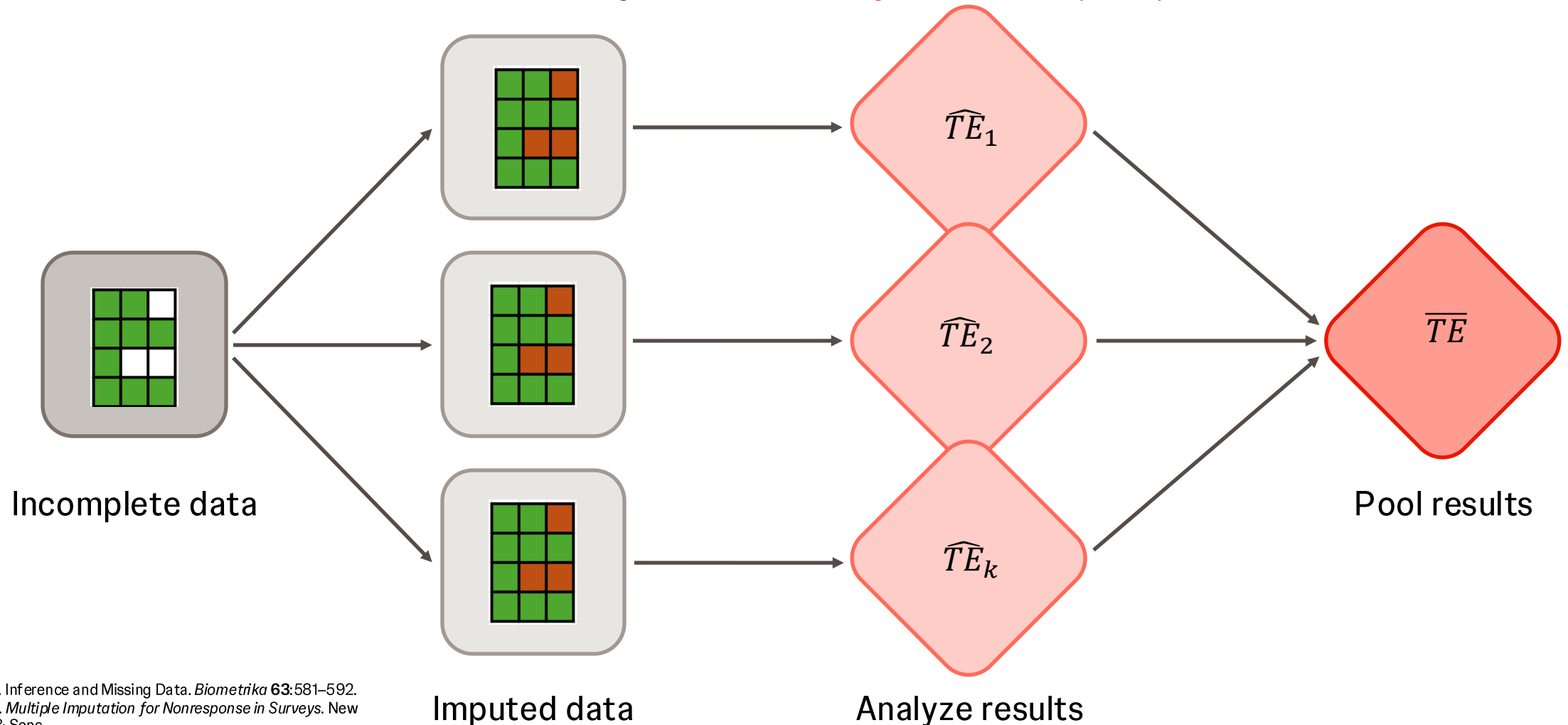
3 additions to CAMIS

1. Tobit regression
2. Reference-based multiple imputation (rbmi)
3. Tipping point analysis

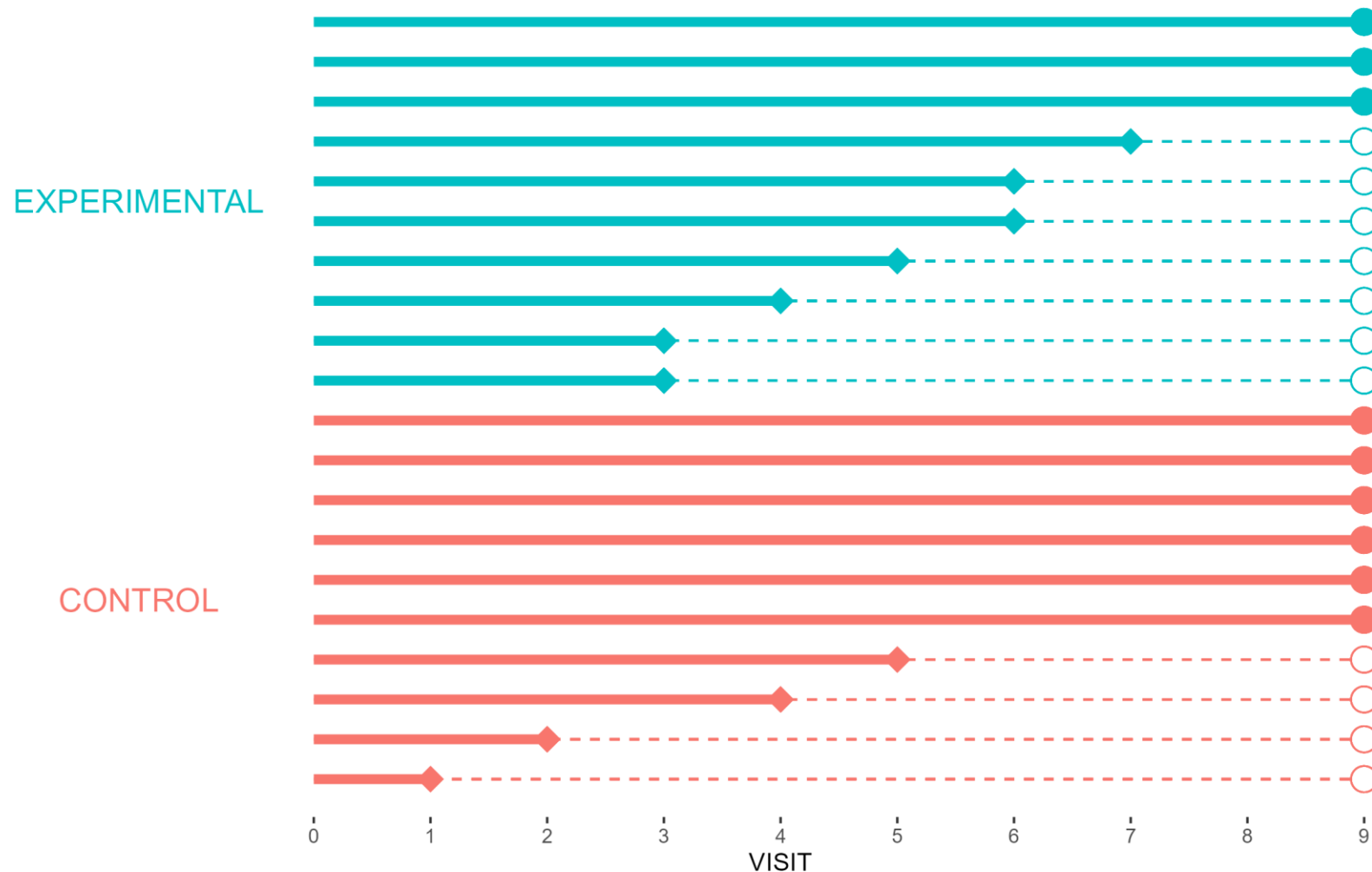


What is (standard) Multiple Imputation?

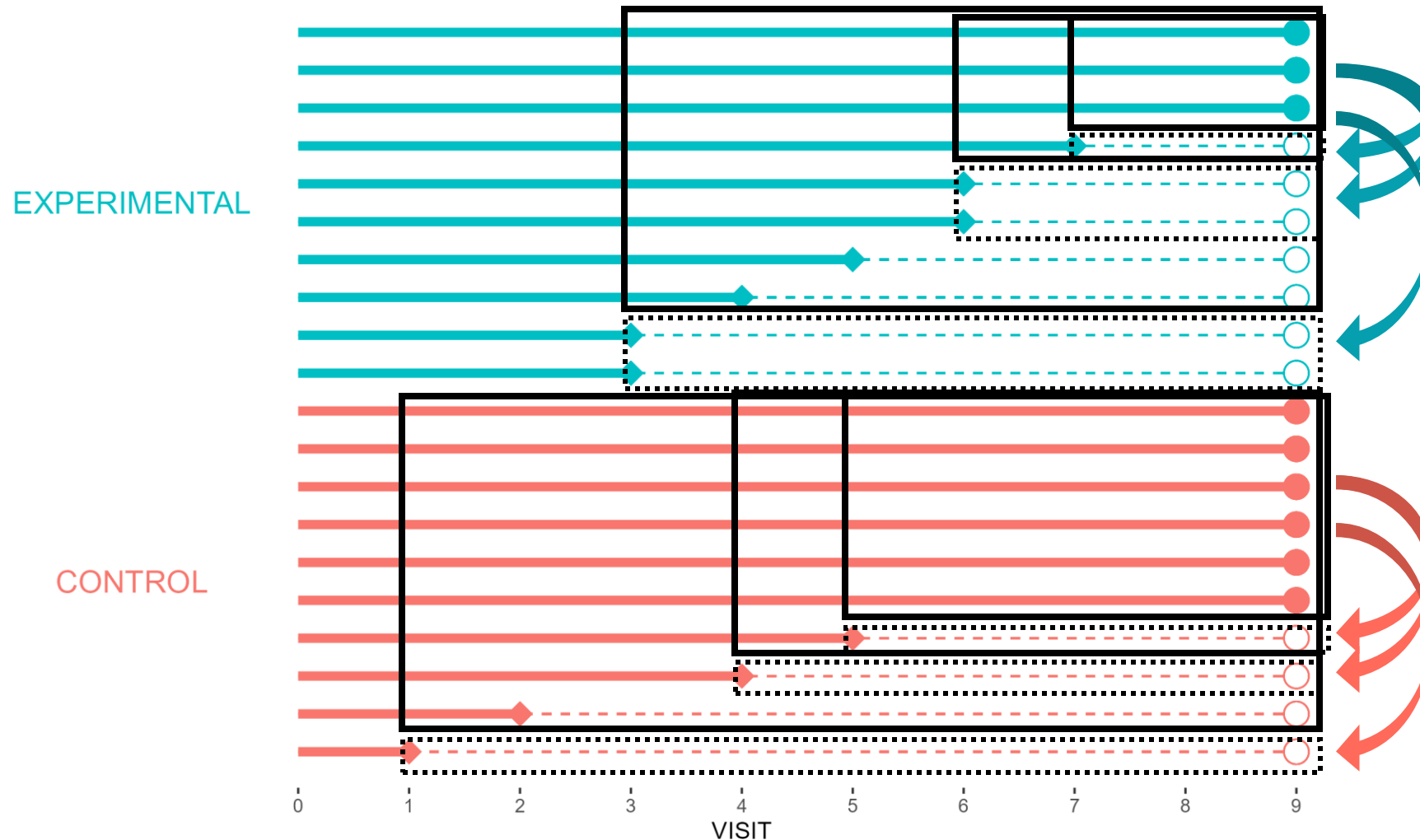
- Theory due to Rubin (1976, 1987)
- Focus is on continuous longitudinal outcomes/endpoints
- Provides valid inference when the missing data are **Missing At Random (MAR)**



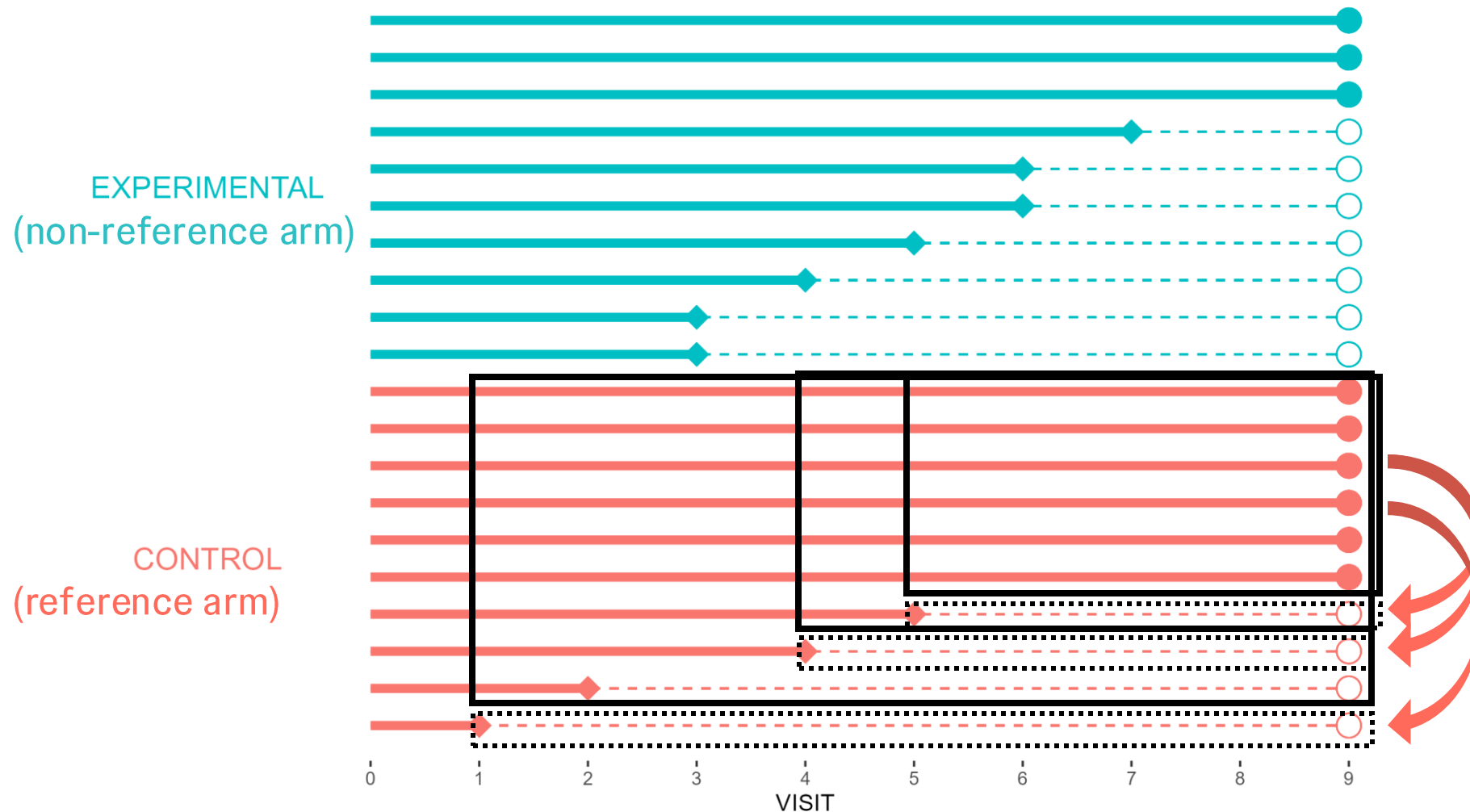
What is (standard) Multiple Imputation?



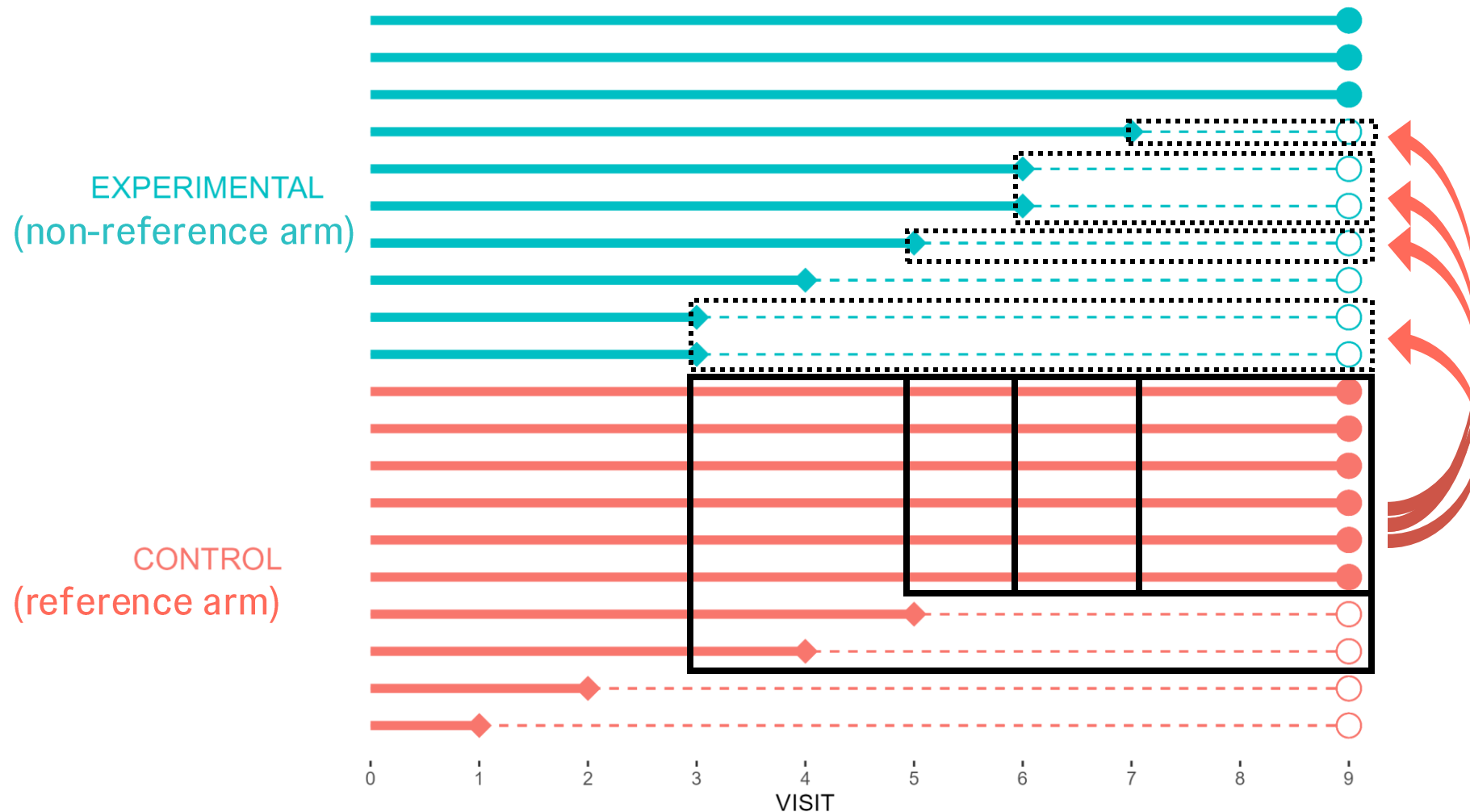
What is (standard) Multiple Imputation?



What is Reference-Based Multiple Imputation?

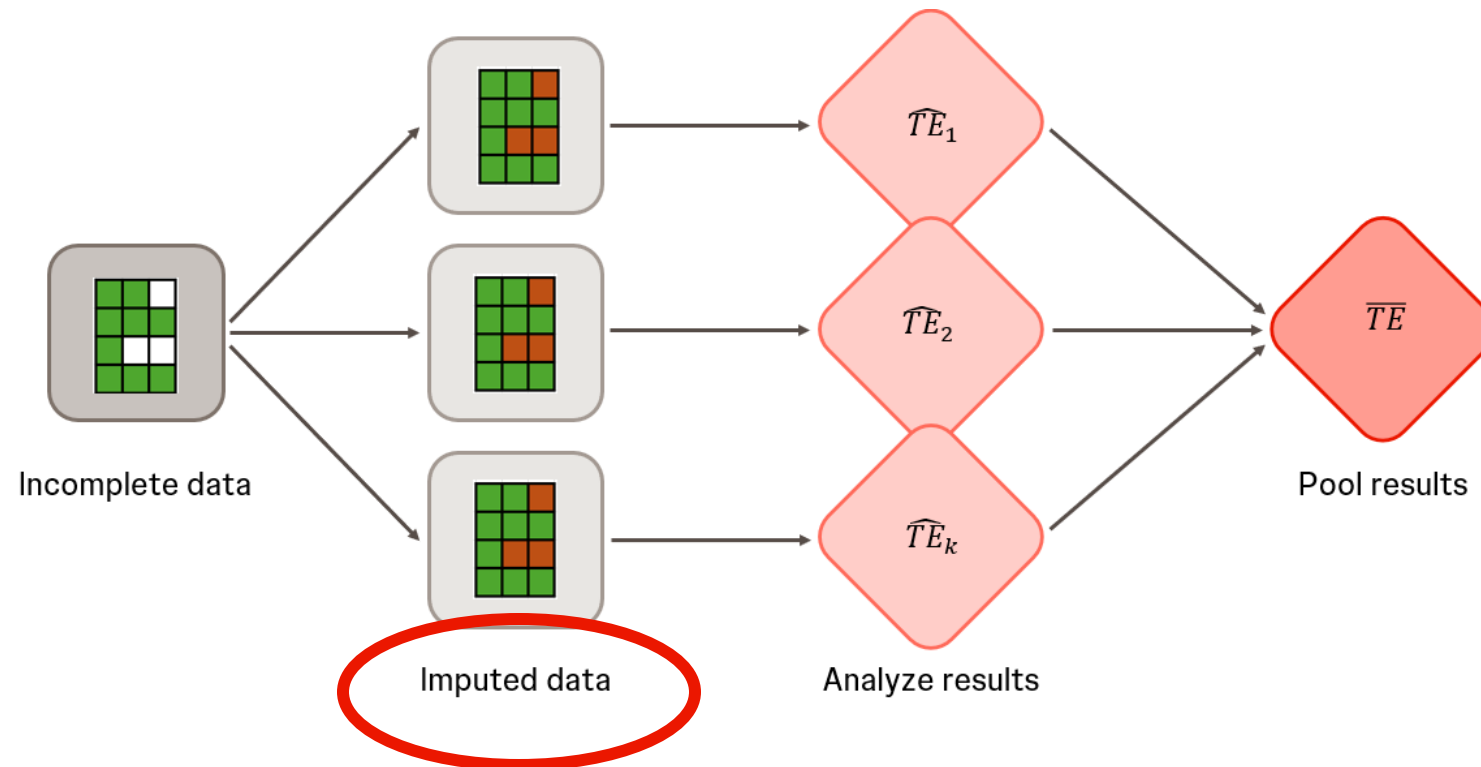


What is Reference-Based Multiple Imputation?



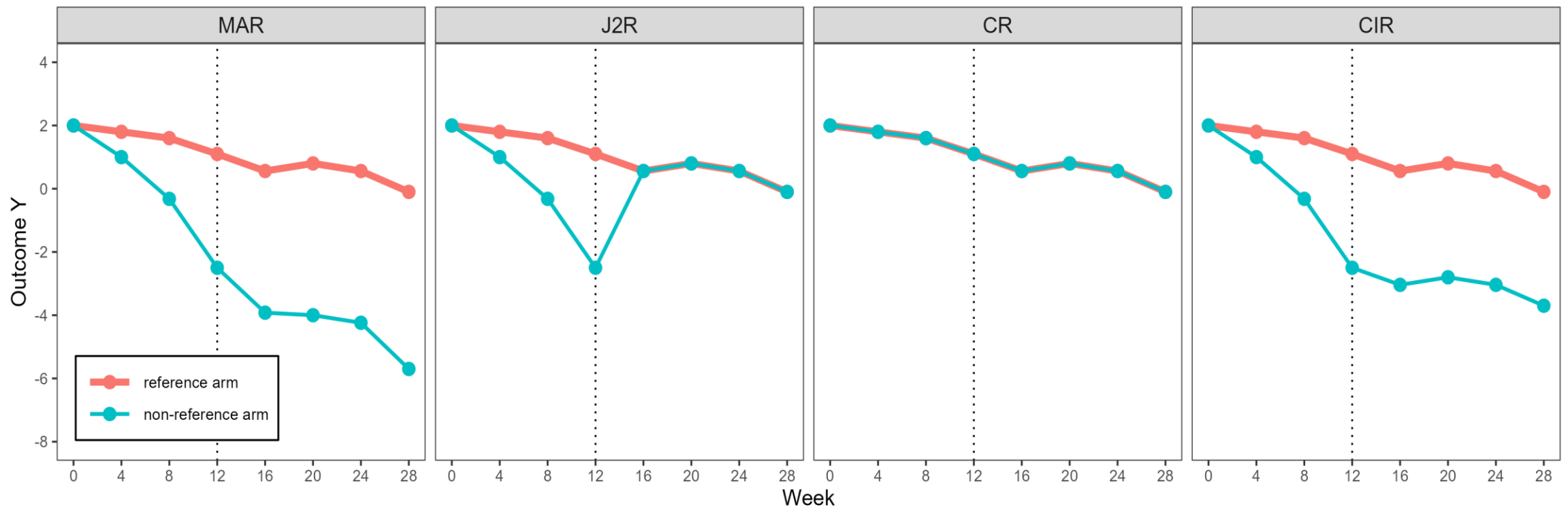
What is Reference-Based Multiple Imputation?

- Theory presented in Carpenter, Roger & Kenward (2013)
- Extends standard MI theory to situations in which **MAR is implausible** (sensitivity analysis)
- Considers **Missing Not At Random (MNAR)** type missingness
- Procedure similar as standard MI, **modifying only the imputation model**



Reference-Based Multiple Imputation strategies

- Considers deviations from MAR, by specifying an imputation model for the unobserved outcomes based on some reference group (e.g., reference arm)
 - ❖ Jump to Reference (J2R)
 - ❖ Copy Reference (CR)
 - ❖ Copy Increments from Reference (CIR)



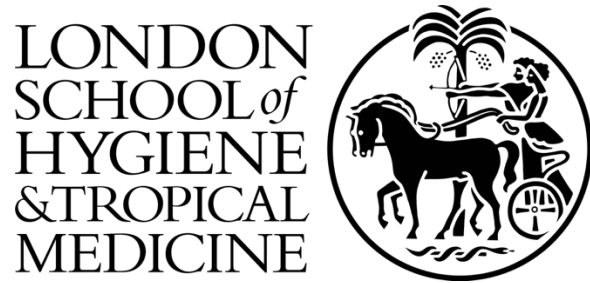
R implementation



Analysis steps

1. fit the imputation model and get Bayesian posterior parameter draws for the parameters
2. use the parameters from the imputation model to generate the imputed datasets
3. run the analysis model on each imputed dataset
4. pool the analysis results across multiple imputed datasets to provide an overall statistic

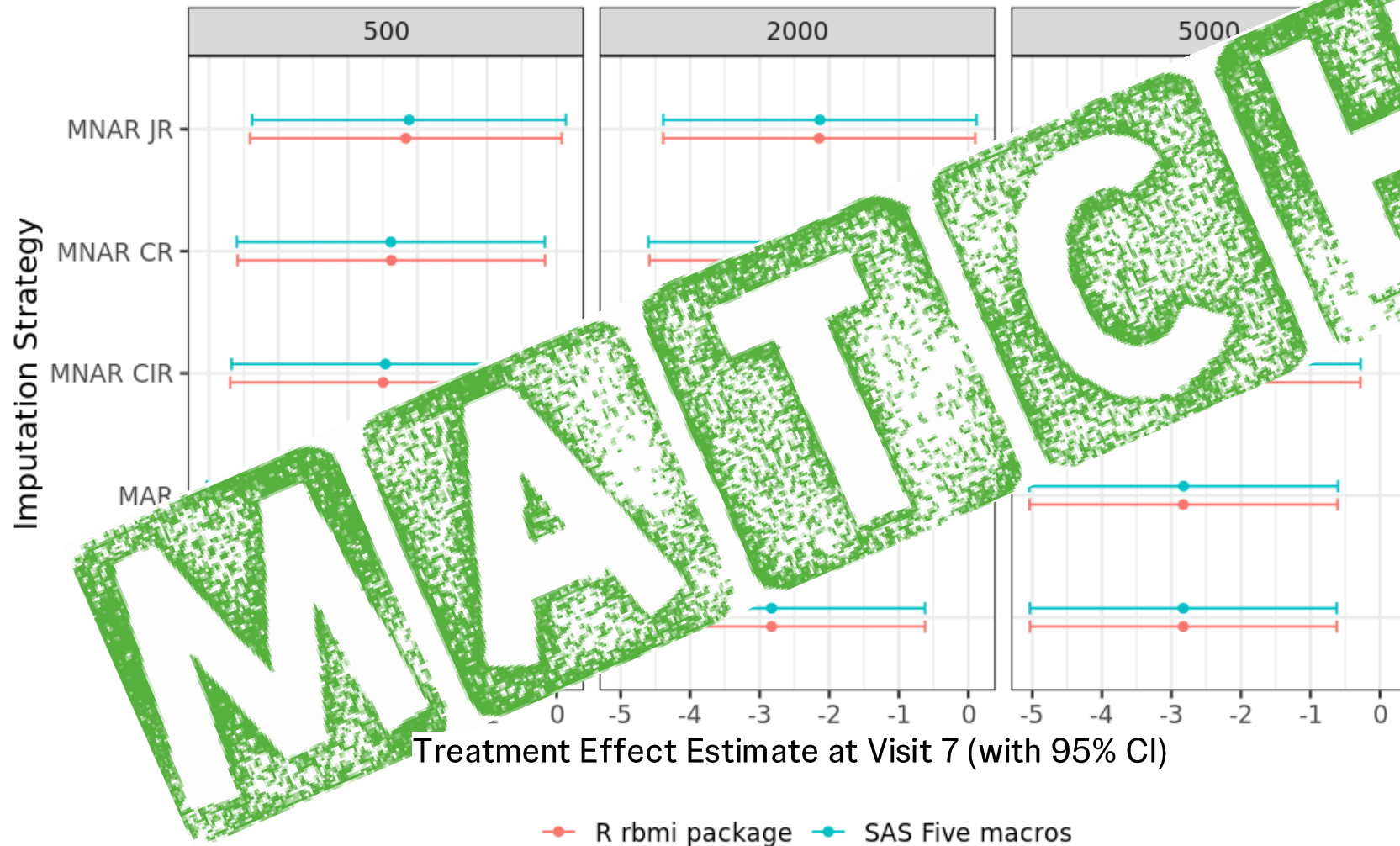
SAS implementation



Five macros

- Part1A declares the parameter estimation model
- Part1B fits the parameter estimation model using MCMC and draws from the joint posterior parameter distribution
- Part2A calculates the predicted mean for each subject
- Part2B imputes missing values
- Part3 carries out the analysis and pools the results

Comparison



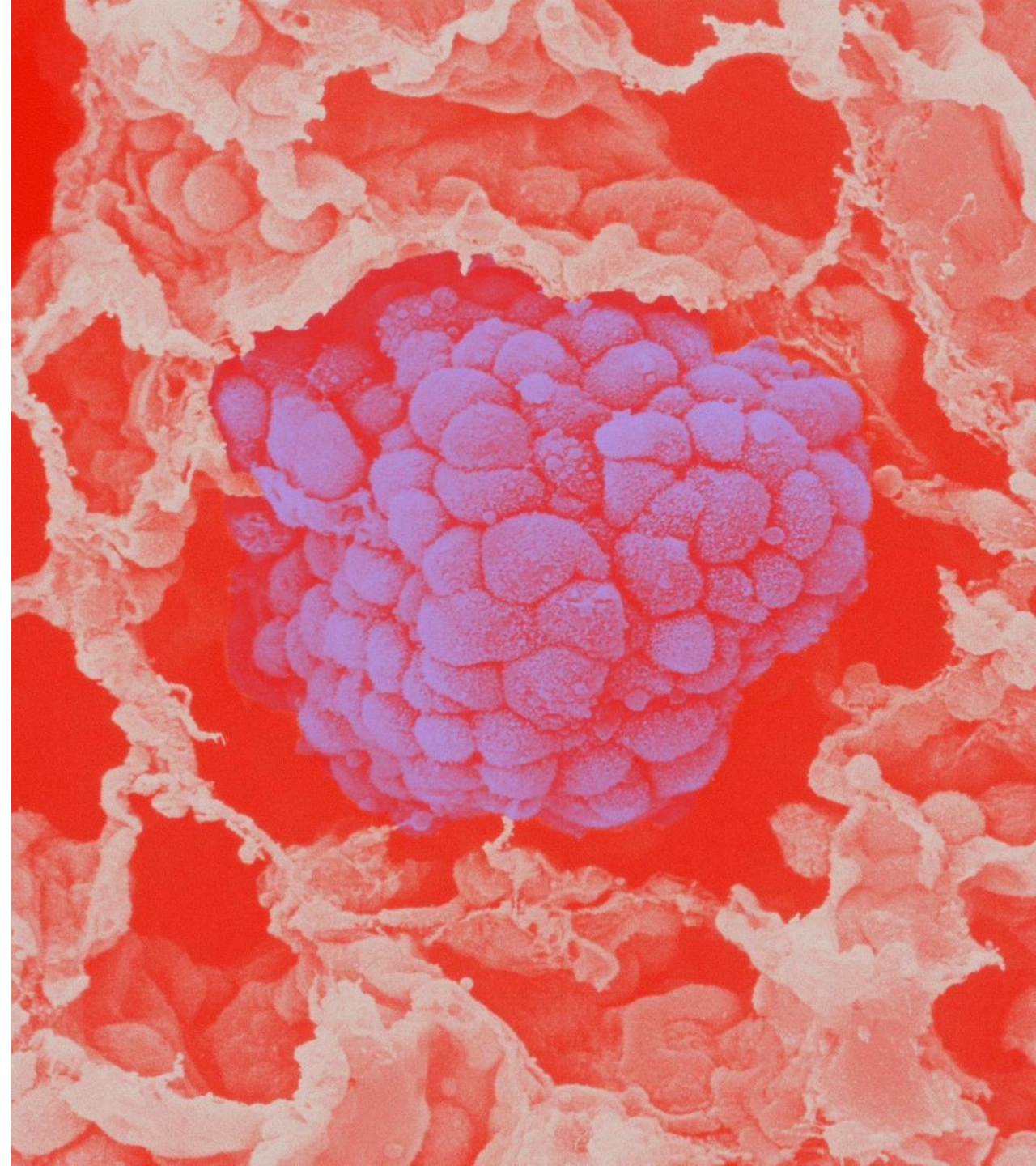
Some Monte simulation error

➤ Five macros faster than **rbmi** (2-3x)

➤ More flexibility in **rbmi** to specify which records need to be imputed

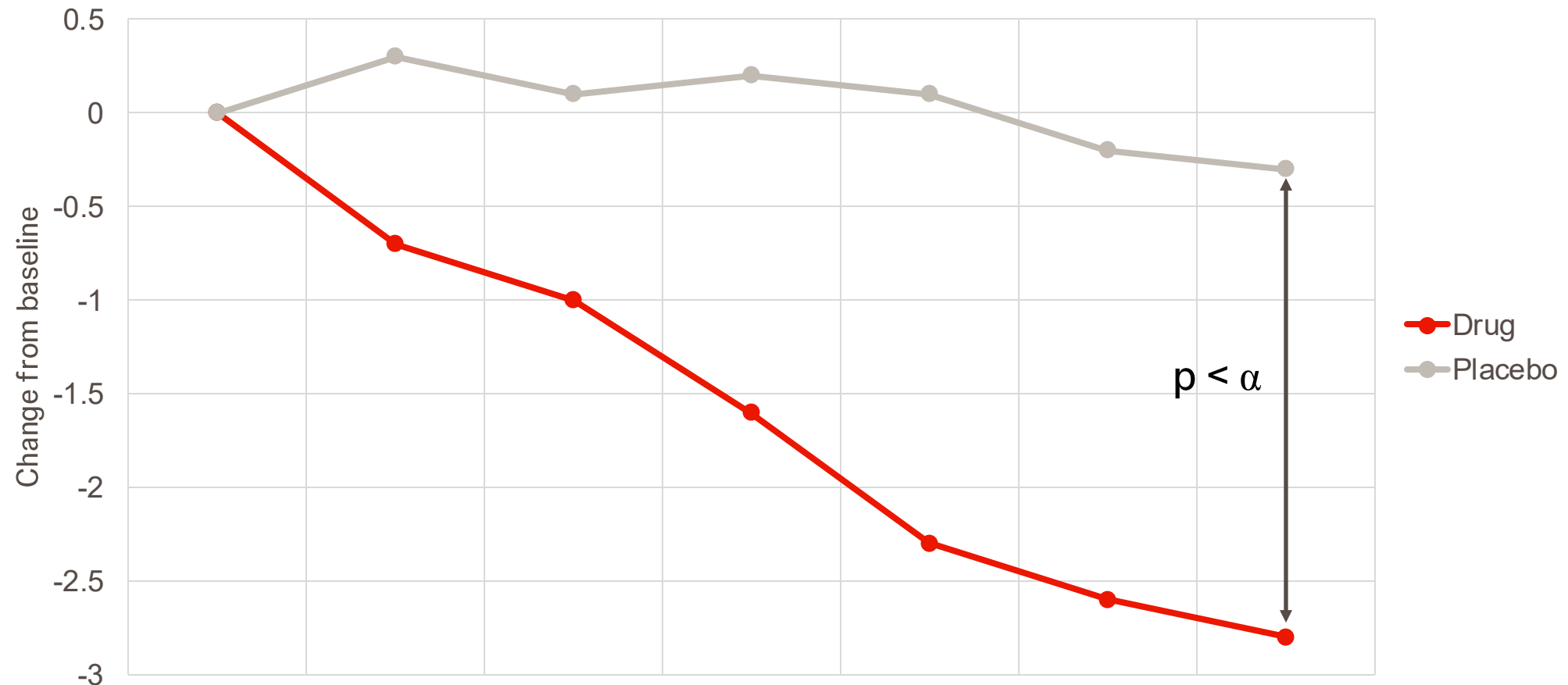
3 additions to CAMIS

1. Tobit regression
2. Reference-based multiple imputation (rbmi)
3. Tipping point analysis



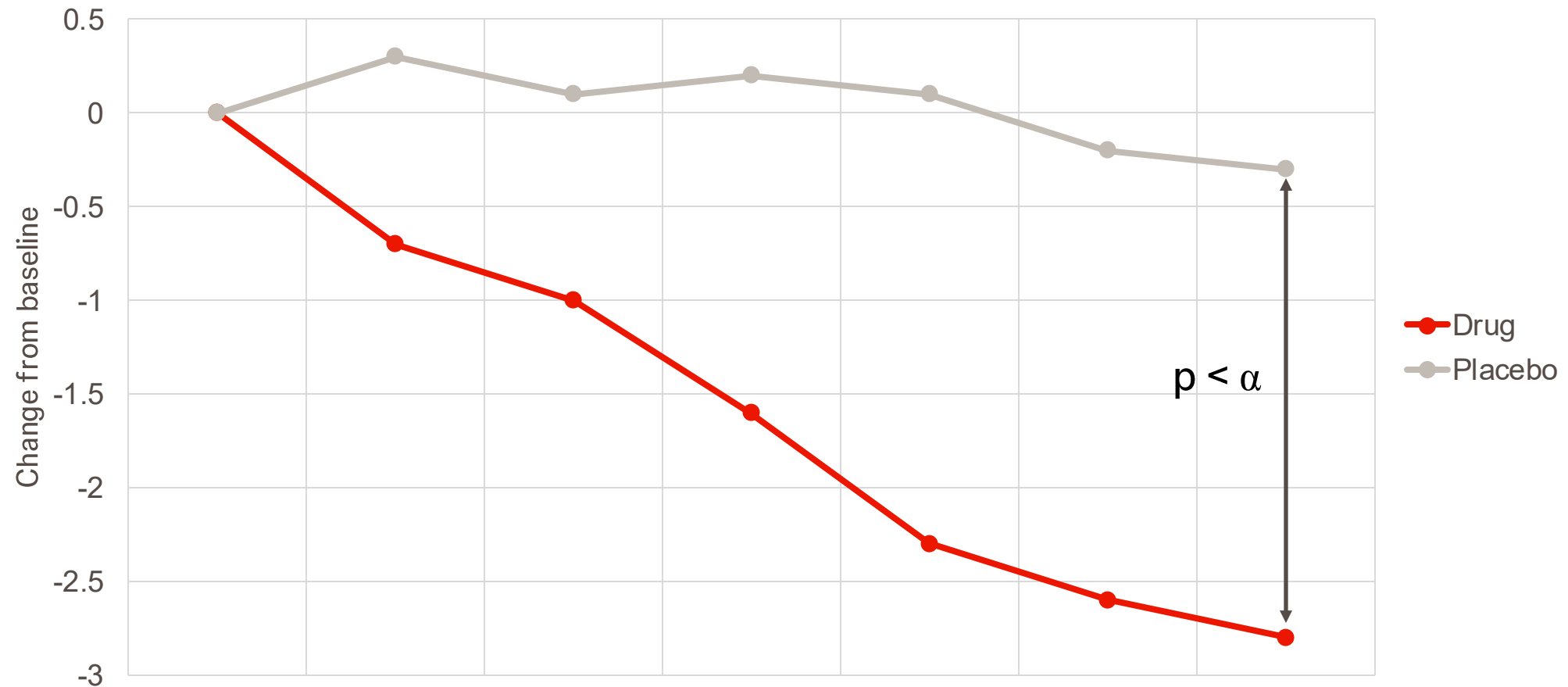
What is Tipping point analysis?

- Continuous longitudinal outcomes/endpoints
- How robust are clinical trial results when they are based on imputed missing data ?
- Sensitivity analysis: Explore influence of missing data on statistical results



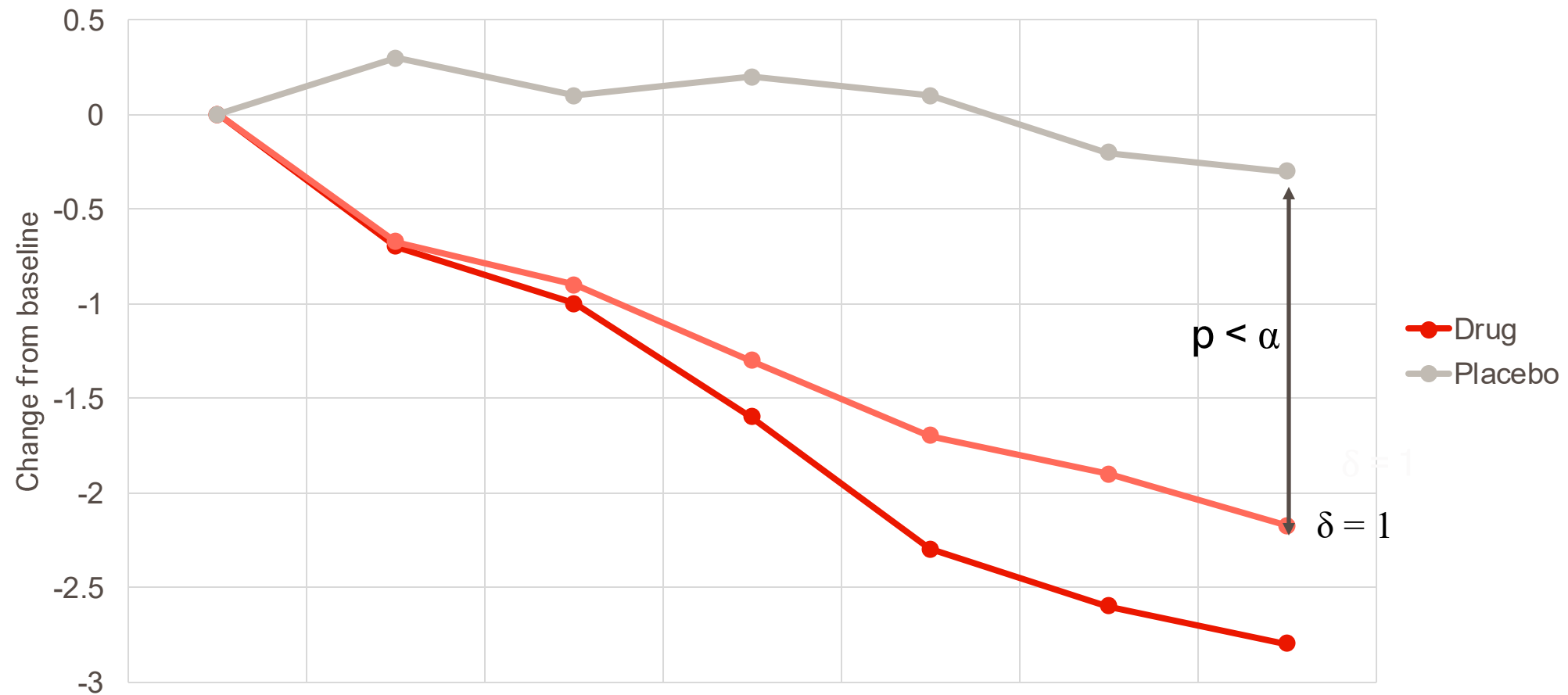
What is Tipping point analysis?

- Shifting imputed missing values in the treatment group towards the reference group until the result becomes non-significant
- Delta adjustment: *imputed missing data point* + δ



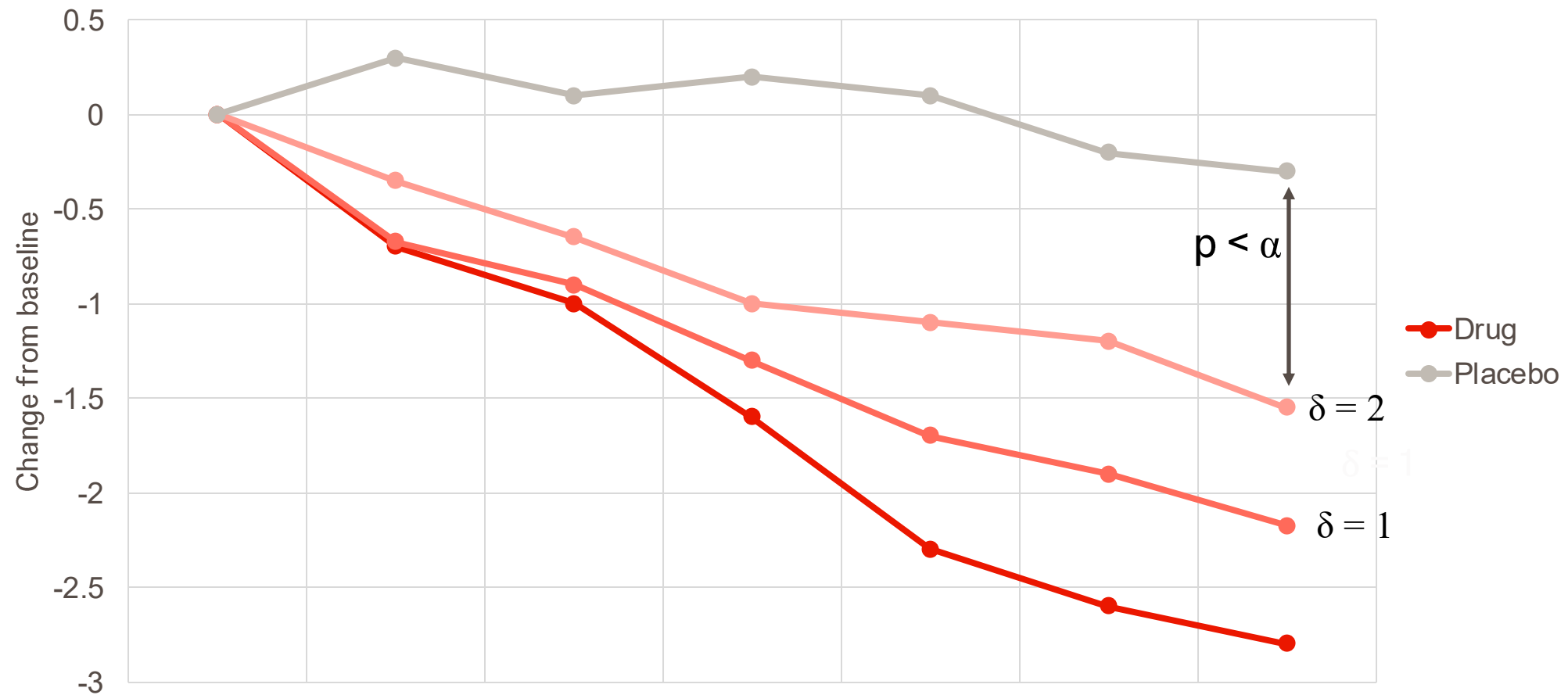
What is Tipping point analysis?

- Shifting imputed missing values in the treatment group towards the reference group until the result becomes non-significant
- Delta adjustment: *imputed missing data point* + δ



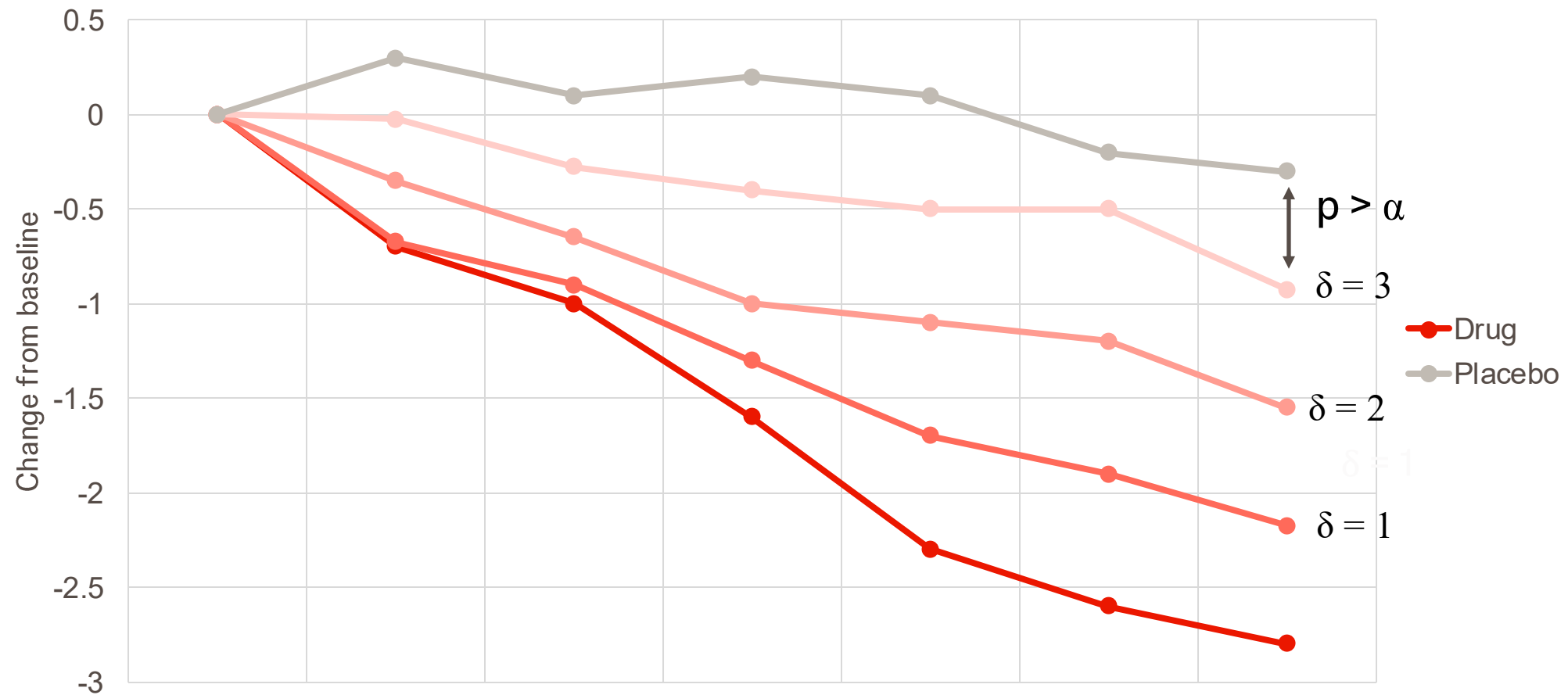
What is Tipping point analysis?

- Shifting imputed missing values in the treatment group towards the reference group until the result becomes non-significant
- Delta adjustment: *imputed missing data point* + δ



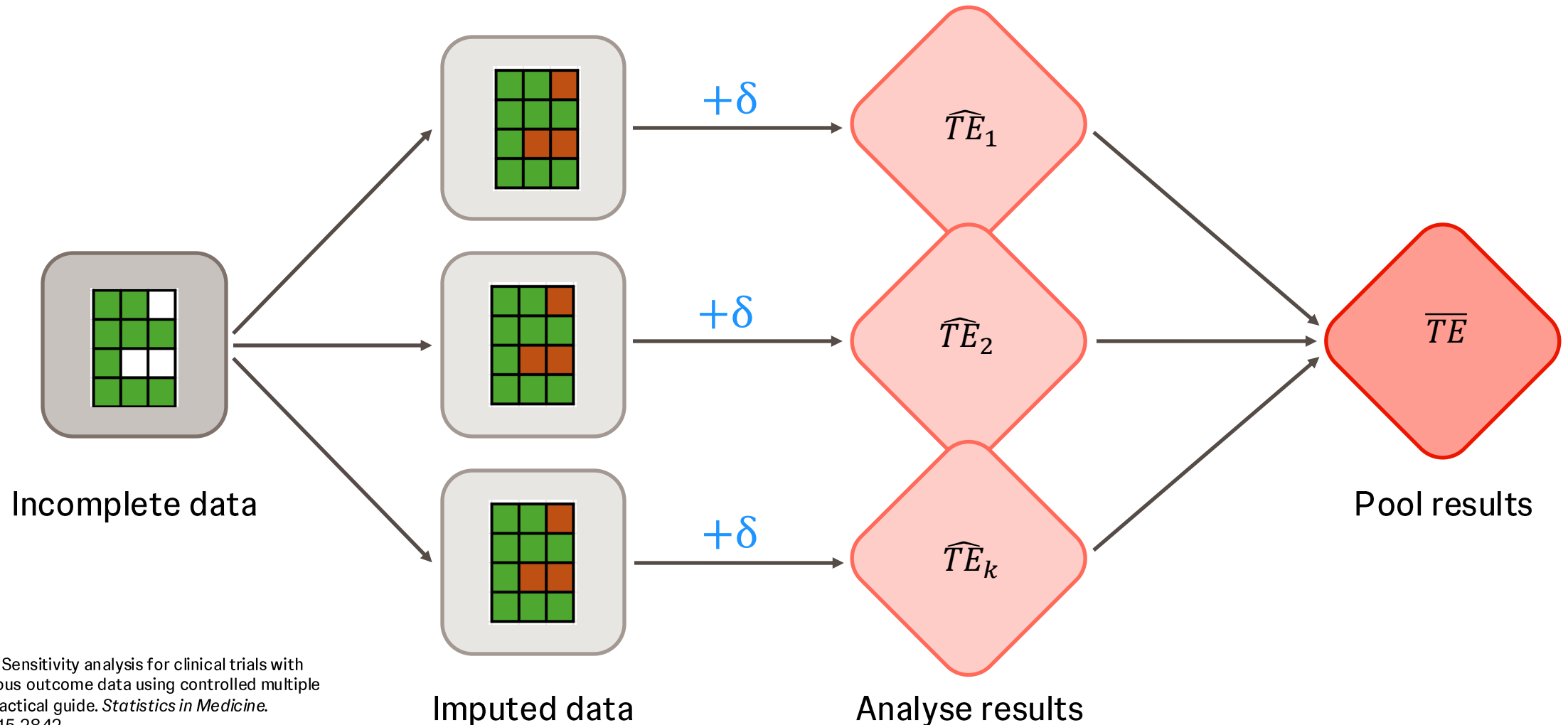
What is Tipping point analysis?

- Shifting imputed missing values in the treatment group towards the reference group until the result becomes non-significant
- Delta adjustment: *imputed missing data point* + δ



Related to Multiple Imputation

- Delta adjustment after data imputation step: *marginal* delta adjustment approach
- Consider a range of delta values



R implementation



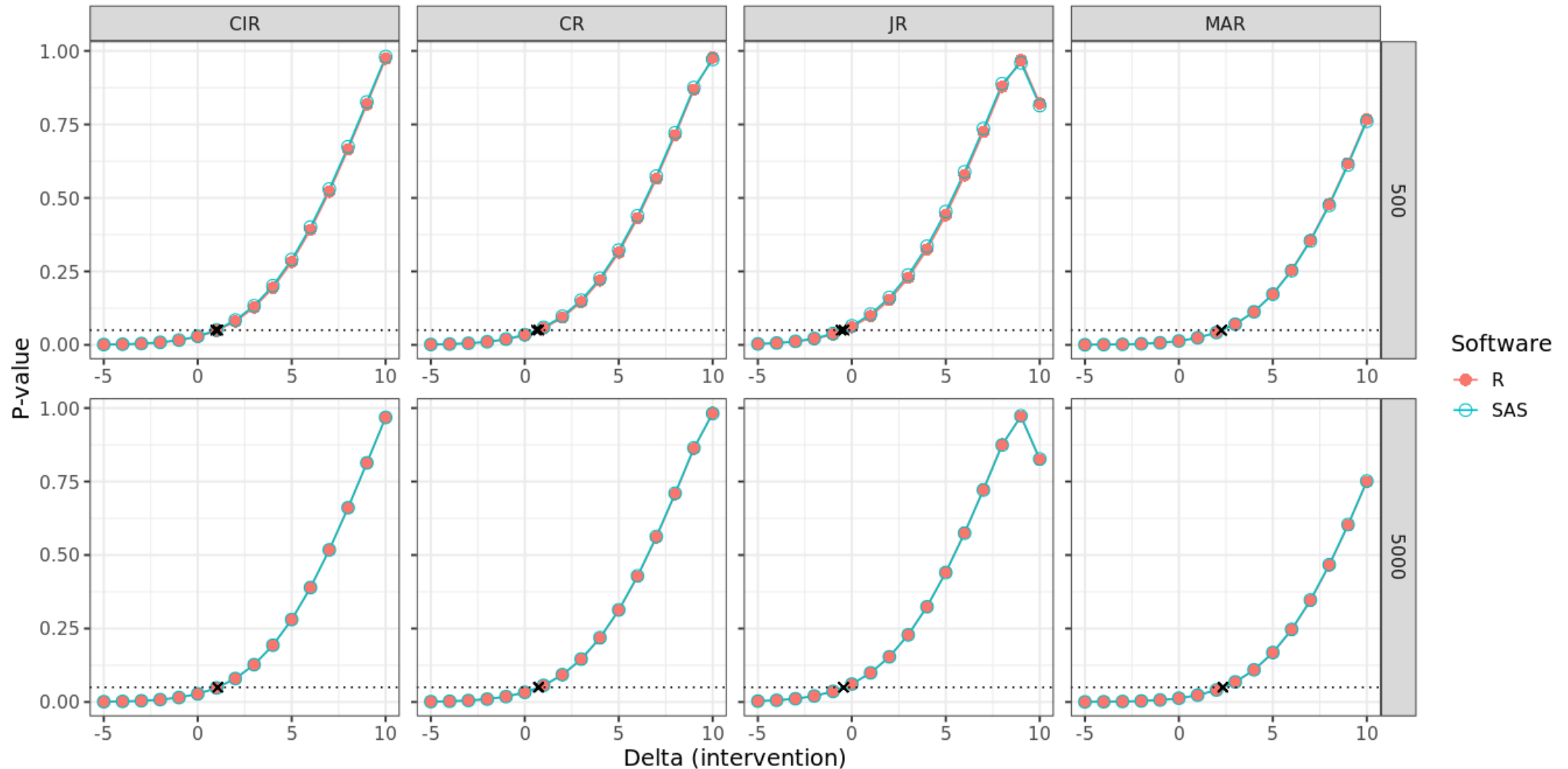
SAS implementation



Five macros

Comparison

- P-value for each delta adjustment (range from -5 to 10)



Comparison

R



- + Perfect match with SAS results
(except for Monte Carlo simulation error)
- + Flexible management of delta adjustments only managed
with standard delta arguments of Part3()
- + Possibility to integrate delta adjustments into
interim analysis
- Slow computation time

SAS Five

- No way of applying delta adjustments to
intermittent missing values
- + Faster computation time

The end? It's only the beginning



- Consult the CAMIS website
- Want to contribute? → Please reach out!
- Community effort



Questions

Comments

Suggestions

