# Graph Based Adaptive MAMS Designs

Cyrus Mehta, Ph.D
Co-Founder and President

Cytel Inc. Cambridge, MA

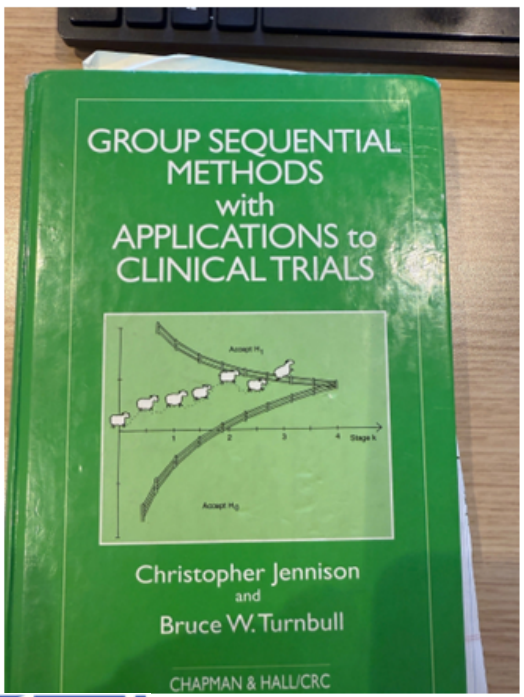
email: mehta@cytel.com – web: www.cytel.com – tel: 617-661-2011

# Acknowledgments


- Joint collaboration with Martin Posch and Ajoy Mukhopadhyay
- Builds upon prior work of:
  - Bretz, Maurer, Brannath, Posch (2009); A graphical approach to sequentially rejective multiple testing procedures
  - Bretz, Posch, Glimm, KlingImueller, Maurer, Rohmeyer (2011); Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes or parametric tests
  - KlingImueller, Posch, Koenig (2014); Adaptive graph-based multiple testing procedures

# History of Adaptive Group Sequential Design

| 1950 to 2000 (Classical) | 2000 to 2025 (Adaptive) |
|---|---|
| • Two arm group sequential designs<br>• Early stopping for efficacy or futility<br>• Flexible monitoring with $\alpha$-spending | **2000-2005 Sample size re-estimation in two-arm designs**<br>• Type-1 error preserved by p-value combination(PVcombo)<br>• Type-1 error preserved conditional error rate (CER)<br>(The two methods are functionally equivalent in 2-arm trials) |




**2005-2010 Seamless phase 2/3 designs**

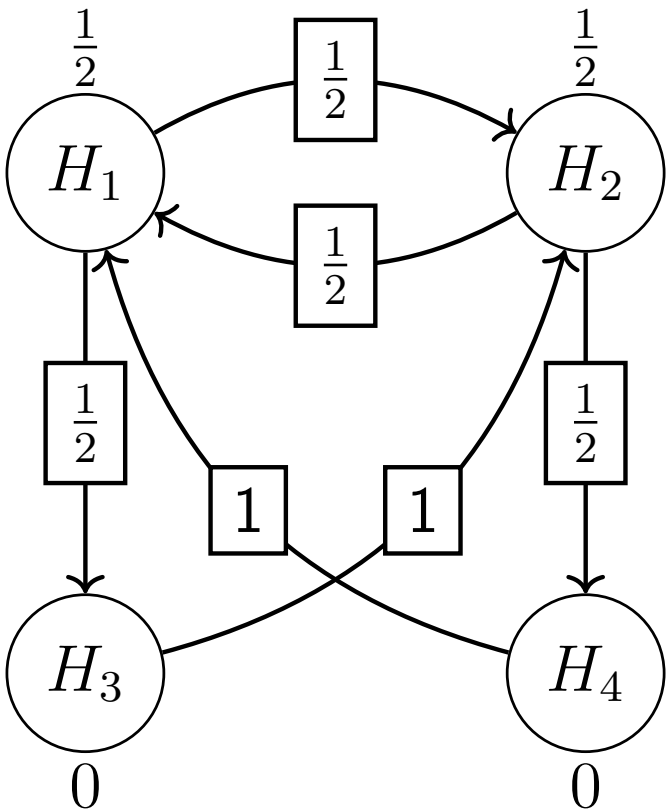
- Test multiple treatments vs control at stage one
- Select a subset of treatments for stage two
- FWER controlled by Closed Testing and PVcombo

**2010-2025 Multi-arm, Multi-stage, Multi-endpoint**

- Extended classical 2-arm designs to MAMS designs
- Adaptive MAMS with SSR and treatment selection
- Inclusion of multiple-endpoints
- Graph based weights assigned to the multiple hypotheses reflecting priorities of study team
- FWER controlled by Closed Testing, Pvcombo, CER

# Schizophrenia Trial: two primary, two secondary endpoints

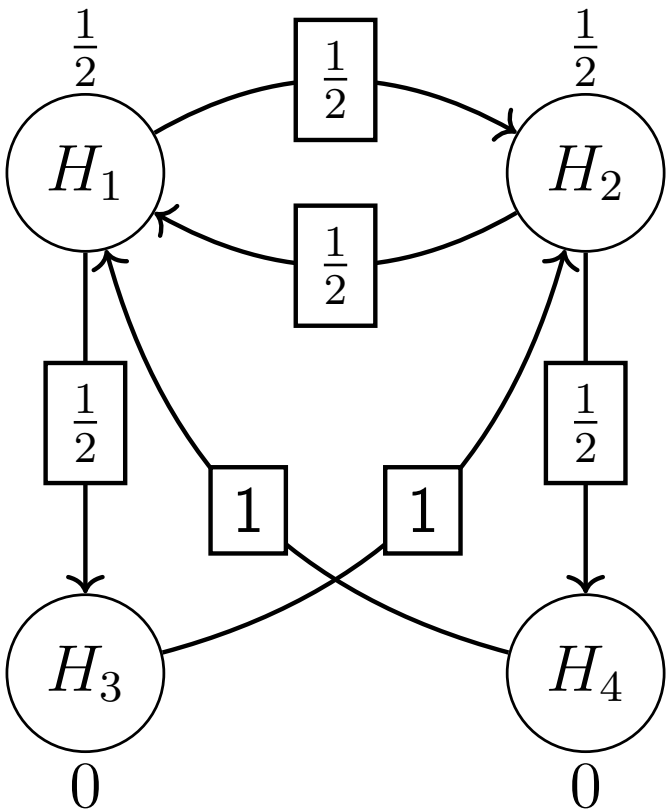
## PANSS score at week 12 for High Dose, Low Dose and Placebo




- H1: Hi vs Pbo; PANSS_tot
- H2: Low vs Pbo; PANSS_tot
- H3: Hi vs Pbo; PANSS_neg
- H4: Low vs Pbo; PANSS_neg

# Schizophrenia Trial: two primary, two secondary endpoints

PANSS score at week 12 for High Dose, Low Dose and Placebo




- H1: Hi vs Pbo; PANSS_tot
- H2: Low vs Pbo; PANSS_tot
- H3: Hi vs Pbo; PANSS_neg
- H4: Low vs Pbo; PANSS_neg

$\text{corr}(H_1, H_2) = 0.5$
$\text{corr}(H_3, H_4) = 0.5$
$\text{corr}(H_1, H_3)$ unknown
$\text{corr}(H_2, H_4)$ unknown

# Sources of Multiplicity and FWER Control

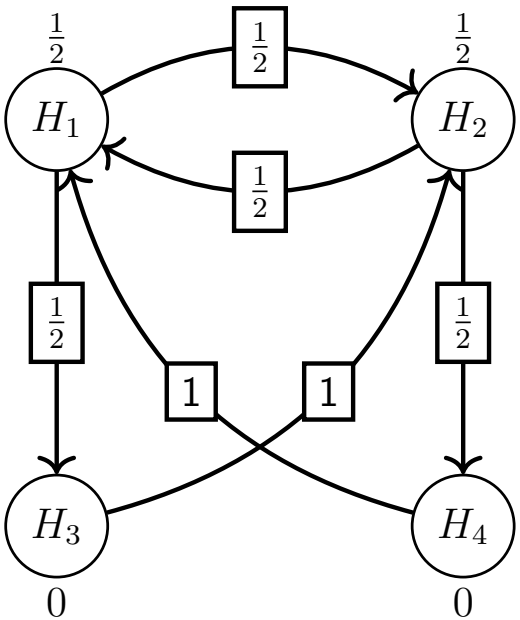
- Can make up to 15 different claims for $\{H_1, H_2, H_3, H_4\}$
- FWER is the probability of making one or more false claims
- Sources of FWER inflation include:
  - Testing multiple hypotheses
  - Testing same hypothesis twice; at stages one and two
  - Making adaptive changes at the end of stage one
- Strong FWER control means FWER $\leq \alpha$ under any configuration of true and false null hypotheses

# Analysis of Schizophrenia Trial

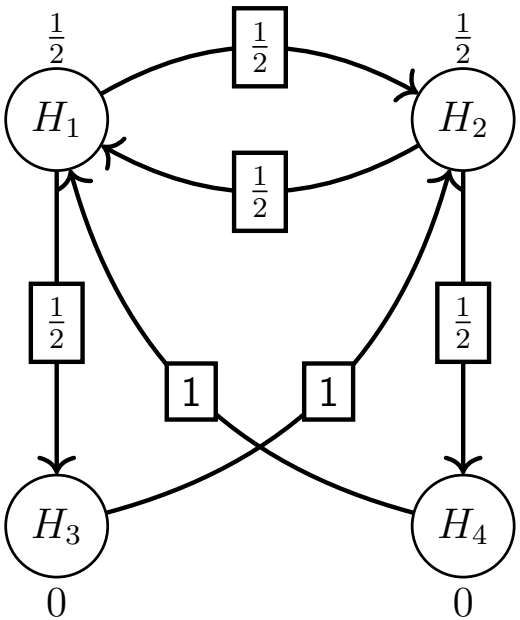
Split $\alpha = 0.025$ into $\alpha_1 = 0.00153$ and $\alpha_2 = 0.02347$
(Lan-DeMets $\alpha$-spending at 50% information)

1. At stage one, test each hypothesis at $\alpha_1 = 0.00153$
2. Eliminate statistically significant hypotheses from further testing
3. Make adaptive decisions for stage two testing
   - drop hypotheses no longer of interest
   - change testing strategy by altering the graph
   - re-assess the sample size and allocation ratio
4. At stage two, test remaining hypotheses with $\alpha_2 = 0.02347$

# Stage One Analysis (at $\alpha = 0.00153$)



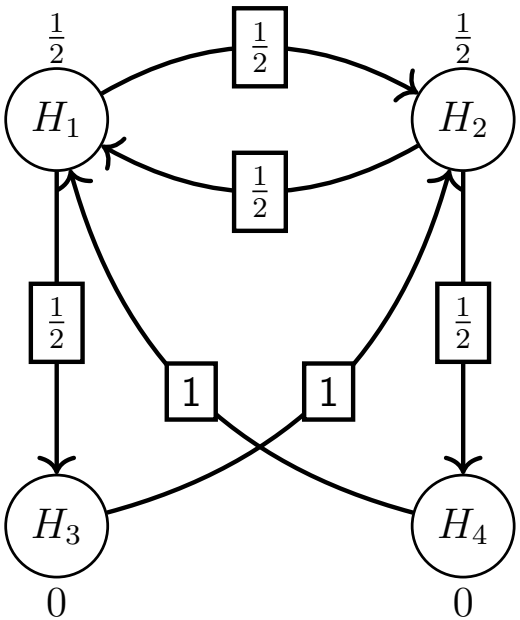
# Stage One Analysis (at $\alpha = 0.00153$)




Decompose graph into 15 weighted intersection hypotheses

| Intersection Hypothesis $H_J$ | Weights $W_J$ | | |
|---|---|---|---|
| $H_1$ | 1 | | |
| $H_2$ | 1 | | |
| $H_3$ | 1 | | |
| $H_4$ | 1 | | |
| $H_1 \cap H_2$ | 0.5, 0.5 | | |
| $H_1 \cap H_3$ | 1.0, 0.0 | | |
| $H_1 \cap H_4$ | 0.75, 0.25 | | |
| $H_2 \cap H_3$ | 0.75, 0.25 | | |
| $H_2 \cap H_4$ | 1.0, 0.0 | | |
| $H_3 \cap H_4$ | 0.5, 0.5 | | |
| $H_1 \cap H_2 \cap H_3$ | 0.5, 0.5, 0.0 | | |
| $H_1 \cap H_2 \cap H_4$ | 0.5, 0.5, 0.0 | | |
| $H_1 \cap H_3 \cap H_4$ | 0.75, 0.0, 0.25 | | |
| $H_2 \cap H_3 \cap H_4$ | 0.75, 0.25, 0.0 | | |
| $H_1 \cap H_2 \cap H_3 \cap H_4$ | 0.5, 0.5, 0.0, 0.0 | | |
| Bretz, Maurer, Brannath, Posch. SiM, 2009 | | | |

# Stage One Analysis (at $\alpha = 0.00153$)




Decompose graph into 15 weighted intersection hypotheses

| Intersection Hypothesis $H_J$ | Weights $W_J$ | | |
|---|---|---|---|
| $H_1$ | 1 | | |
| $H_2$ | 1 | | |
| $H_3$ | 1 | | |
| $H_4$ | 1 | | |
| $H_1 \cap H_2$ | 0.5, 0.5 | | |
| $H_1 \cap H_3$ | 1.0, 0.0 | | |
| $H_1 \cap H_4$ | 0.75, 0.25 | | |
| $H_2 \cap H_3$ | 0.75, 0.25 | | |
| $H_2 \cap H_4$ | 1.0, 0.0 | | |
| $H_3 \cap H_4$ | 0.5, 0.5 | | |
| $H_1 \cap H_2 \cap H_3$ | 0.5, 0.5, 0.0 | | |
| $H_1 \cap H_2 \cap H_4$ | 0.5, 0.5, 0.0 | | |
| $H_1 \cap H_3 \cap H_4$ | 0.75, 0.0, 0.25 | | |
| $H_2 \cap H_3 \cap H_4$ | 0.75, 0.25, 0.0 | | |
| $H_1 \cap H_2 \cap H_3 \cap H_4$ | 0.5, 0.5, 0.0, 0.0 | | |
| Bretz, Maurer, Brannath, Posch. SiM, 2009 | | | |

- Unadjusted (raw) Stage One p-values $= \{0.00045, 0.0952, 0.0225, 0.1104\}$

# Aside: Computing Multiplicity Adjusted P-values

$$\{p_1 = 0.00045, p_2 = 0.0952, p_3 = 0.0225, p_4 = 0.1104\}$$

- Use raw p-values as building blocks for adjusted p-values
- Adjustment may be parametric or nonparametric

# Aside: Computing Multiplicity Adjusted P-values

$$\{p_1 = 0.00045, p_2 = 0.0952, p_3 = 0.0225, p_4 = 0.1104\}$$

- Use raw p-values as building blocks for adjusted p-values
- Adjustment may be parametric or nonparametric

Nonparametric (Bonferroni): Test $H_2 \cap H_3$ with weights $w_2 = 0.75, w_3 = 0.25$

$$p_{\{1,2\}} = \min\{\frac{p_2}{w_2}, \frac{p_3}{w_3}\} = \min\{\frac{0.0952}{0.75}, \frac{0.0225}{0.25}\} = 0.09$$

# Aside: Computing Multiplicity Adjusted P-values

$$\{p_1 = 0.00045, p_2 = 0.0952, p_3 = 0.0225, p_4 = 0.1104\}$$

- Use raw p-values as building blocks for adjusted p-values
- Adjustment may be parametric or nonparametric

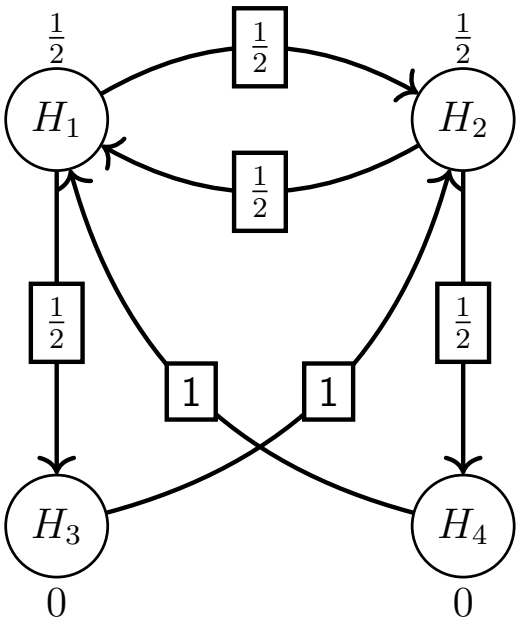
Nonparametric (Bonferroni): Test $H_2 \cap H_3$ with weights $w_2 = 0.75, w_3 = 0.25$

$$p_{\{1,2\}} = \min\{\frac{p_2}{w_2}, \frac{p_3}{w_3}\} = \min\{\frac{0.0952}{0.75}, \frac{0.0225}{0.25}\} = 0.09$$

Parametric (Dunnett): Test $H_1 \cap H_2$ with weights $w_1 = 0.5, w_2 = 0.5$

$$p_{\{2,3\}} = \Pr\{\min[\frac{P_1}{w_1}, \frac{P_2}{w_2}] \leq \min[\frac{0.00045}{0.5}, \frac{0.0952}{0.5}]\} = 0.00088$$

# Stage One Analysis (at $\alpha = 0.00153$)

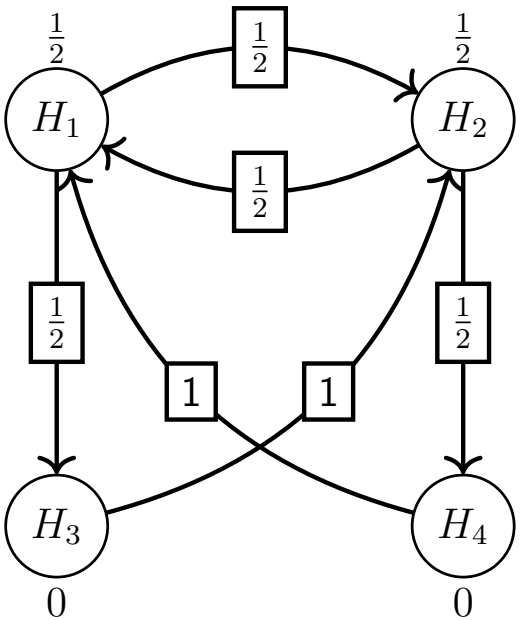



Decompose graph into 15 weighted intersection hypotheses

| Intersection Hypothesis $H_J$ | Weights $W_J$ | Adjusted p-value $p_{J,1}$ | Adjustment Method |
|---|---|---|---|
| $H_1$ | 1 | 0.00045 | NA |
| $H_2$ | 1 | 0.0952 | NA |
| $H_3$ | 1 | 0.0225 | NA |
| $H_4$ | 1 | 0.1104 | NA |
| $H_1 \cap H_2$ | 0.5, 0.5 | 0.00088 | Parametric |
| $H_1 \cap H_3$ | 1.0, 0.0 | 0.00045 | Bonferroni |
| $H_1 \cap H_4$ | 0.75, 0.25 | 0.0006 | Bonferroni |
| $H_2 \cap H_3$ | 0.75, 0.25 | 0.0900 | Bonferroni |
| $H_2 \cap H_4$ | 1.0, 0.0 | 0.0952 | Bonferroni |
| $H_3 \cap H_4$ | 0.5, 0.5 | 0.0410 | Parametric |
| $H_1 \cap H_2 \cap H_3$ | 0.5, 0.5, 0.0 | 0.00088 | Mixed |
| $H_1 \cap H_2 \cap H_4$ | 0.5, 0.5, 0.0 | 0.00088 | Mixed |
| $H_1 \cap H_3 \cap H_4$ | 0.75, 0.0, 0.25 | 0.0006 | Mixed |
| $H_2 \cap H_3 \cap H_4$ | 0.75, 0.25, 0.0 | 0.0900 | Mixed |
| $H_1 \cap H_2 \cap H_3 \cap H_4$ | 0.5, 0.5, 0.0, 0.0 | 0.00088 | Mixed |
| Bretz, Maurer, Brannath, Posch. SiM, 2009 | | | |

- Unadjusted (raw) Stage One p-values $= \{0.00045, 0.0952, 0.0225, 0.1104\}$

# Stage One Analysis (at $\alpha = 0.00153$)

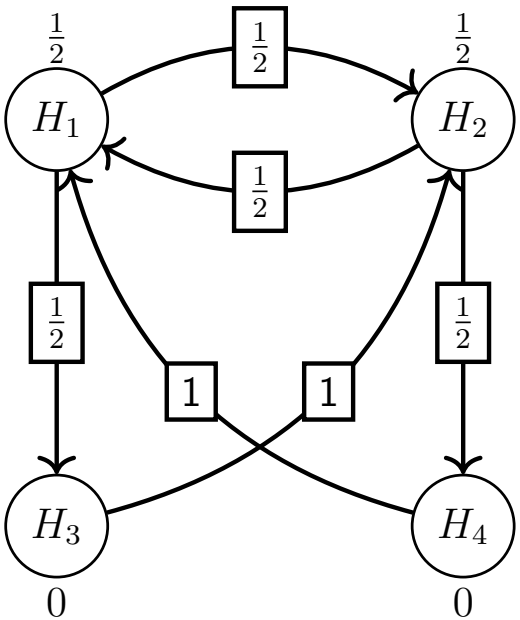



Decompose graph into 15 weighted intersection hypotheses

| Intersection Hypothesis $H_J$ | Weights $W_J$ | Adjusted p-value $p_{J,1}$ | Adjustment Method |
|---|---|---|---|
| $H_1$ | 1 | 0.00045 | NA |
| $H_2$ | 1 | 0.0952 | NA |
| $H_3$ | 1 | 0.0225 | NA |
| $H_4$ | 1 | 0.1104 | NA |
| $H_1 \cap H_2$ | 0.5, 0.5 | 0.00088 | Parametric |
| $H_1 \cap H_3$ | 1.0, 0.0 | 0.00045 | Bonferroni |
| $H_1 \cap H_4$ | 0.75, 0.25 | 0.0006 | Bonferroni |
| $H_2 \cap H_3$ | 0.75, 0.25 | 0.0900 | Bonferroni |
| $H_2 \cap H_4$ | 1.0, 0.0 | 0.0952 | Bonferroni |
| $H_3 \cap H_4$ | 0.5, 0.5 | 0.0410 | Parametric |
| $H_1 \cap H_2 \cap H_3$ | 0.5, 0.5, 0.0 | 0.00088 | Mixed |
| $H_1 \cap H_2 \cap H_4$ | 0.5, 0.5, 0.0 | 0.00088 | Mixed |
| $H_1 \cap H_3 \cap H_4$ | 0.75, 0.0, 0.25 | 0.0006 | Mixed |
| $H_2 \cap H_3 \cap H_4$ | 0.75, 0.25, 0.0 | 0.0900 | Mixed |
| $H_1 \cap H_2 \cap H_3 \cap H_4$ | 0.5, 0.5, 0.0, 0.0 | 0.00088 | Mixed |
| Bretz, Maurer, Brannath, Posch. SiM, 2009 | | | |

- Unadjusted (raw) Stage One p-values $= \{0.00045, 0.0952, 0.0225, 0.1104\}$
- Reject any $H_J$ for which $p_{J,1} \leq 0.00153$ (Lan–DeMets)

# Stage One Analysis (at $\alpha = 0.00153$)

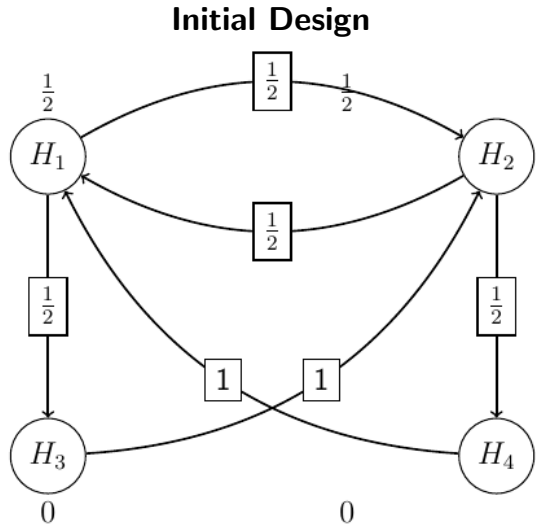



Decompose graph into 15 weighted intersection hypotheses

| Intersection Hypothesis $H_J$ | Weights $W_J$ | Adjusted p-value $p_{J,1}$ | Adjustment Method |
|---|---|---|---|
| $H_1$ | 1 | 0.00045 | NA |
| $H_2$ | 1 | 0.0952 | NA |
| $H_3$ | 1 | 0.0225 | NA |
| $H_4$ | 1 | 0.1104 | NA |
| $H_1 \cap H_2$ | 0.5, 0.5 | 0.00088 | Parametric |
| $H_1 \cap H_3$ | 1.0, 0.0 | 0.00045 | Bonferroni |
| $H_1 \cap H_4$ | 0.75, 0.25 | 0.0006 | Bonferroni |
| $H_2 \cap H_3$ | 0.75, 0.25 | 0.0900 | Bonferroni |
| $H_2 \cap H_4$ | 1.0, 0.0 | 0.0952 | Bonferroni |
| $H_3 \cap H_4$ | 0.5, 0.5 | 0.0410 | Parametric |
| $H_1 \cap H_2 \cap H_3$ | 0.5, 0.5, 0.0 | 0.00088 | Mixed |
| $H_1 \cap H_2 \cap H_4$ | 0.5, 0.5, 0.0 | 0.00088 | Mixed |
| $H_1 \cap H_3 \cap H_4$ | 0.75, 0.0, 0.25 | 0.0006 | Mixed |
| $H_2 \cap H_3 \cap H_4$ | 0.75, 0.25, 0.0 | 0.0900 | Mixed |
| $H_1 \cap H_2 \cap H_3 \cap H_4$ | 0.5, 0.5, 0.0, 0.0 | 0.00088 | Mixed |
| Bretz, Maurer, Brannath, Posch. SiM, 2009 | | | |

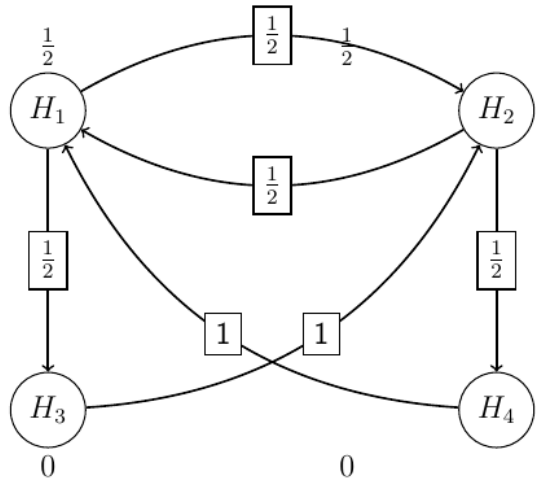
- Unadjusted (raw) Stage One p-values $= \{0.00045, 0.0952, 0.0225, 0.1104\}$
- Reject any $H_J$ for which $p_{J,1} \leq 0.00153$ (Lan–DeMets)
- $H_1$ is rejected under closed testing. $H_2, H_3, H_4$ are retained

# Modify Graph as Desired for Stage Two

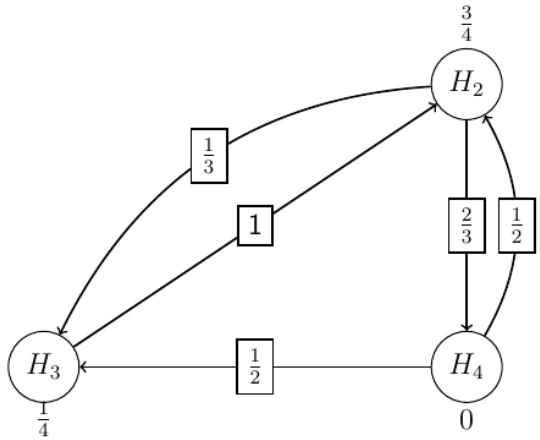


# Modify Graph as Desired for Stage Two

**Initial Design**

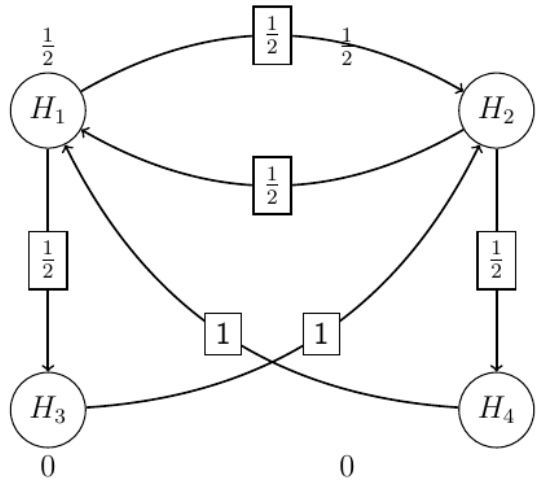



**$H_1$ rejected at interim analysis**

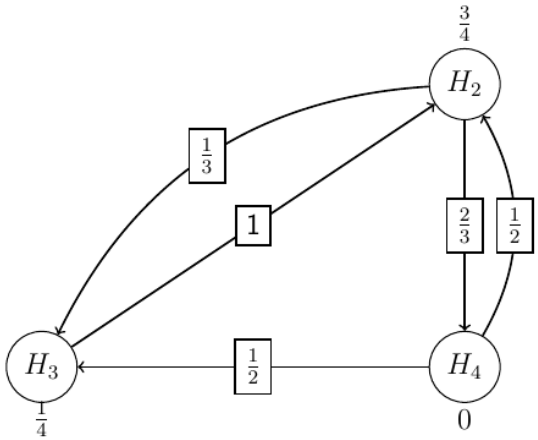


# Modify Graph as Desired for Stage Two

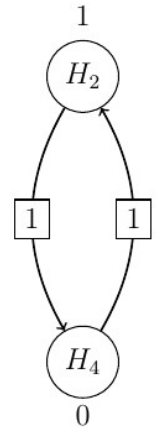
**Initial Design**




**$H_1$ rejected at interim analysis**

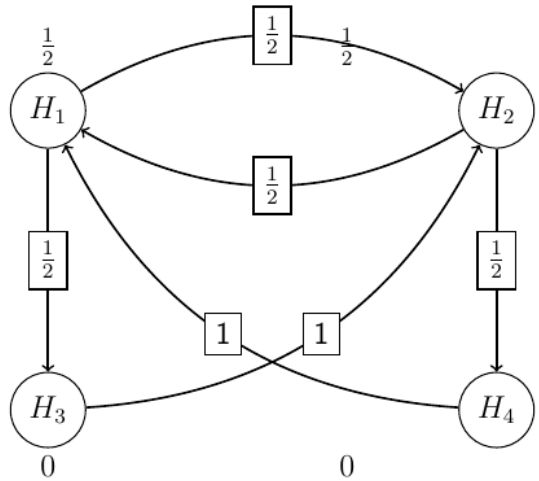



**$H_1$ rejected, $H_3$ dropped**

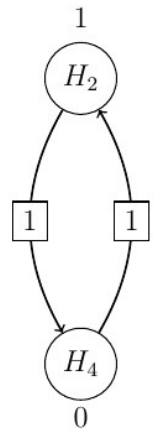


# Modify Graph as Desired for Stage Two

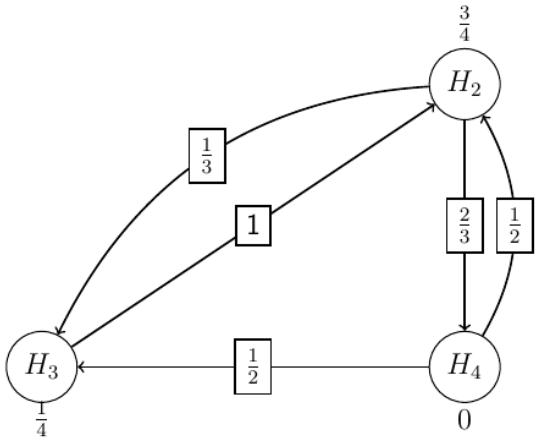
**Initial Design**



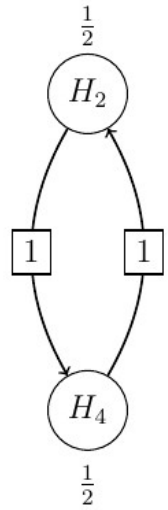

**$H_1$ rejected, $H_3$ dropped**




**$H_1$ rejected at interim analysis**




**$H_1$ rejected, $H_3$ dropped, weights altered**



# Two Methods for Stage Two Testing

## P-value Combination Method (PVcombo)

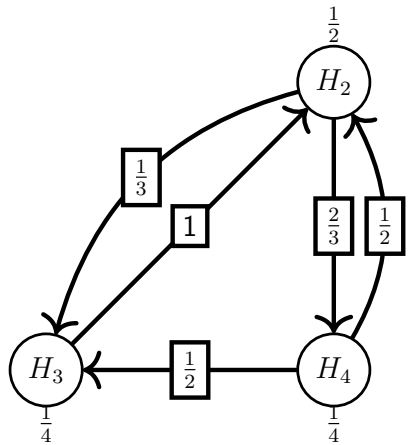
- Combine incremental multiplicity adjusted p-values
- Pre-specify the combination function
- Use two-arm group sequential cut-off to reject hypotheses
- Use closed testing of individual hypotheses

## Conditional Error Rate Method (CER)

- Use unadjusted cumulative p-values from the two stages
- Use multi-arm group sequential cut-off to reject hypotheses
- Adjust cut-off by preserving CER if design is adapted

# Stage Two Testing by PVcombo Method

- Suppose we proceed with a modified stage two graph




- Decompose into 7 weighted intersection hypothesis tests

| Index Set $J \in \mathcal{J}^+$ | Intersection Hypothesis $H_J$ | Weights $W_J$ | Adjusted p-value $p_{J,(2)}$ (incremental) | Adjustment Method |
|---|---|---|---|---|
| {2} | $H_2$ | 1 | 0.1121 | NA |
| {3} | $H_3$ | 1 | 0.0112 | NA |
| {4} | $H_4$ | 1 | 0.0448 | NA |
| {2, 3} | $H_2 \cap H_3$ | 0.625, 0.375 | 0.0299 | Bonferroni |
| {2, 4} | $H_2 \cap H_4$ | 0.75, 0.25 | 0.1495 | Bonferroni |
| {3, 4} | $H_3 \cap H_4$ | 0.417, 0.583 | 0.0249 | Parametric |
| {2, 3, 4} | $H_2 \cap H_3 \cap H_4$ | 0.5, 0.25, 0.25 | 0.0418 | Mixed |

# Stage Two Testing by PVcombo Method

- Combine incremental stagewise adjusted p-values $p_{J,1}$ and $p_{J,(2)}$
- Utilize a pre-specified combination function

$$C(p_{J,1}, p_{J,(2)}) = 1 - \Phi[\sqrt{\frac{1}{2}}\Phi^{-1}(1 - p_{J,1}) + \sqrt{\frac{1}{2}}\Phi^{-1}(1 - p_{J,(2)})]$$

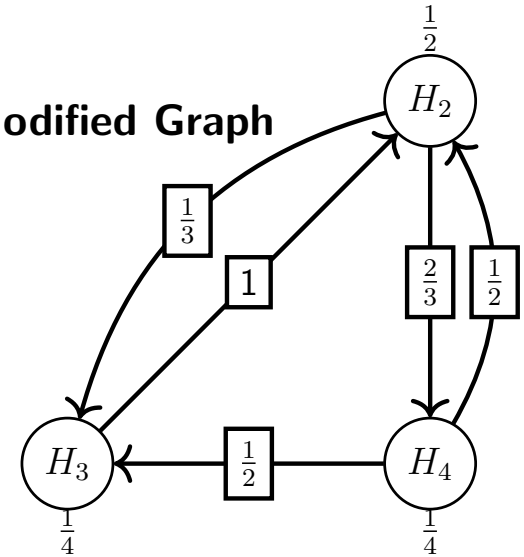
- Reject $H_J$ if $C(p_{J,1}, p_{J,(2)}) \leq 0.0245$ ( Lan-DeMets Boundary)

| Index Set $J \in \mathcal{J}^+$ | Intersection Hypothesis $H_J$ | Stage-wise Adjusted p-values | | Combined p-value $C_J(p_{J,1}, p_{J.(2)})$ |
|---|---|---|---|---|
| | | $p_{J,1}$ | $p_{J,(2)}$ | |
| {2} | $H_2$ | 0.0952 | 0.1121 | 0.0371 |
| {3} | $H_3$ | 0.0225 | 0.0112 | 0.0012 |
| {4} | $H_4$ | 0.1104 | 0.0448 | 0.0194 |
| {2, 3} | $H_2 \cap H_3$ | 0.0900 | 0.0299 | 0.0113 |
| {2, 4} | $H_2 \cap H_4$ | 0.0952 | 0.1495 | 0.0484 |
| {3, 4} | $H_3 \cap H_4$ | 0.0410 | 0.0249 | 0.0044 |
| {2, 3, 4} | $H_2 \cap H_3 \cap H_4$ | 0.0900 | 0.0418 | 0.0149 |

Reject $H_3$ by the closed test

# Stage Two Testing by CER Method

**Modified Graph**




| Intersection $H_J$ | Modified Weights $W_J$ | Multiplicity Adjustment |
|---|---|---|
| $H_2$ | 1 | NA |
| $H_3$ | 1 | NA |
| $H_4$ | 1 | NA |
| $H_2 \cap H_3$ | 0.625, 0.375 | Nonparametric |
| $H_2 \cap H_4$ | 0.75, 0.25 | Nonparametric |
| $H_3 \cap H_4$ | 0.417, 0.583 | Parametric |
| $H_2 \cap H_3 \cap H_4$ | 0.5, 0.25, 0.25 | Mixed |

## Example: Stage Two Adaptive Test of $H_3 \cap H_4$

- Pre-specify level-$\alpha$ multivariate group sequential boundaries $(C_1, C_2)$:
  $$P\left\{\min_{\{j\in(3,4)\}}[p_{j,1} \le w_j C_1] \text{ or } \min_{\{j\in(3,4)\}}[p_{j,2} \le w_j C_2]\right\} = \alpha$$
- If trial adapts, recompute the stage two cut-off $C_2$ such that
  CER of adaptive test $\le$ CER of pre-specified test

# Stage Two Testing by CER Method

Example: Stage Two Adaptive Test of $H_3 \cap H_4$

- Notation:
  $(p_{3,1}, p_{4,1})$: stage one unadjusted p-values
  $(p_{3,2}, p_{4,2})$: stage two cumulative, unadjusted p-values if no adaptation
  $(p_{3,2}*, p_{4,2}*)$: stage twocumulative, unadjusted p-values if trial adapts

# Stage Two Testing by CER Method

Example: Stage Two Adaptive Test of $H_3 \cap H_4$

- Notation:
  $(p_{3,1}, p_{4,1})$: stage one unadjusted p-values
  $(p_{3,2}, p_{4,2})$: stage two cumulative, unadjusted p-values if no adaptation
  $(p_{3,2}*, p_{4,2}*)$: stage twocumulative, unadjusted p-values if trial adapts

- If no adaptation. Reject $H_3 \cap H_4$ at stage two if

$$\min\left(\frac{p_{3,2}}{w_3}, \frac{p_{4,2}}{w_4}\right) \leq C_2$$

  where $C_2$ is the group sequential cut-off for 2-arms vs control

# Stage Two Testing by CER Method

Example: Stage Two Adaptive Test of $H_3 \cap H_4$

- Notation:
  $(p_{3,1}, p_{4,1})$: stage one unadjusted p-values
  $(p_{3,2}, p_{4,2})$: stage two cumulative, unadjusted p-values if no adaptation
  $(p_{3,2}*, p_{4,2}*)$: stage twocumulative, unadjusted p-values if trial adapts

- If no adaptation. Reject $H_3 \cap H_4$ at stage two if

$$\min\left(\frac{p_{3,2}}{w_3}, \frac{p_{4,2}}{w_4}\right) \leq C_2$$

  where $C_2$ is the group sequential cut-off for 2-arms vs control

- If adaptation. Reject $H_3 \cap H_4$ if

$$\min\left(\frac{p^*_{3,2}}{w^*_3}, \frac{p^*_{4,2}}{w_4}^*\right) \leq C^*_2$$

# Stage Two Testing by CER Method

$C_2^*$ satisfies the CER condition

$$P_0\{\min(\frac{P_{3,2}^*}{w_3^*}, \frac{P_{4,2}^*}{w_4}^*) \leq C_2^* | p_{3,1}, p_{4,1}\}$$

$$\leq P_0\{\min(\frac{P_{3,2}}{w_3}, \frac{P_{4,2}}{w_4}) \leq C_2 | p_{3,1}, p_{4,1}\}$$

## Key Idea of the CER Method

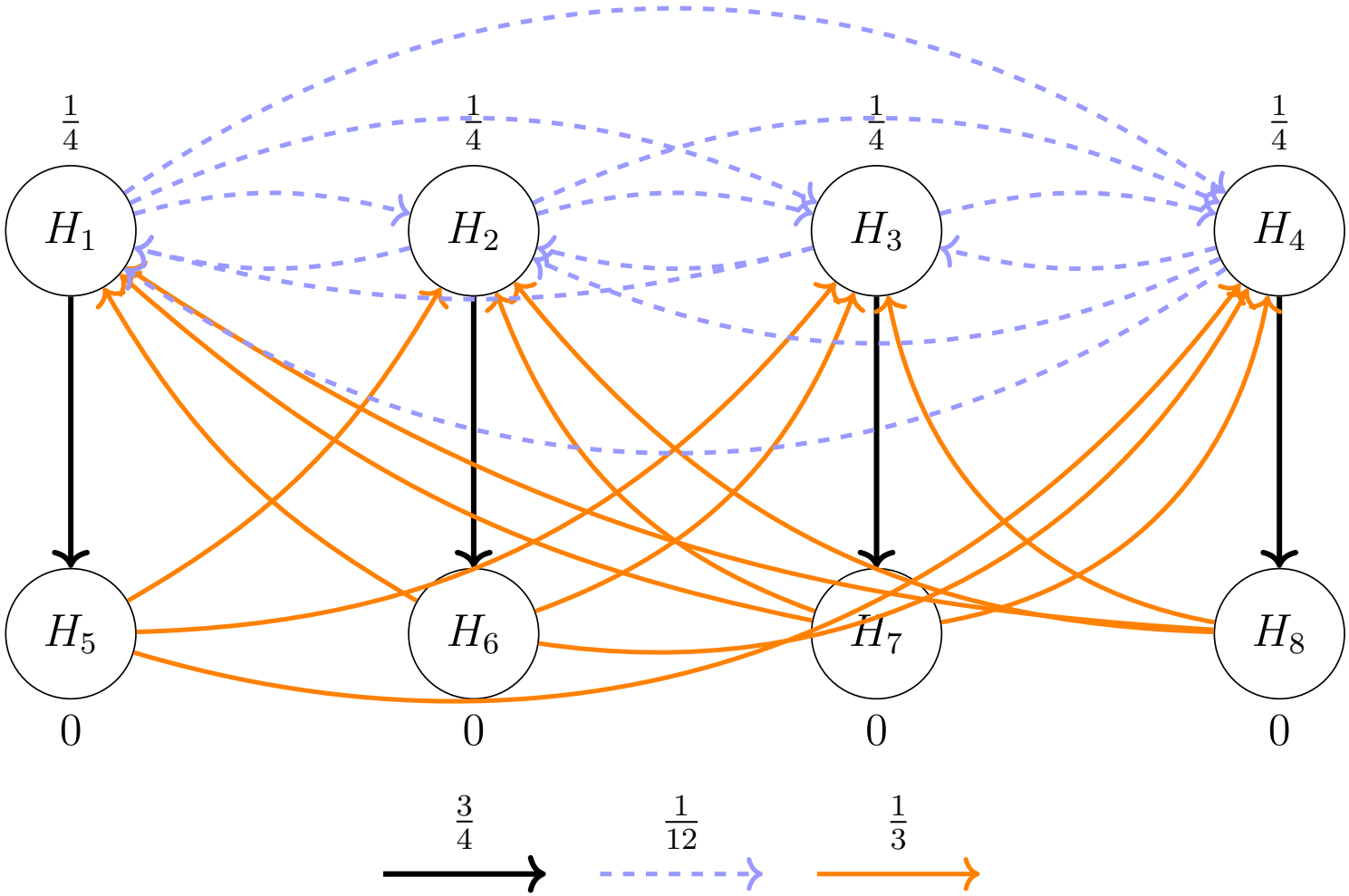
A procedure that bounds the conditional type-1 error of the adapted test, by the conditional type-1 error of the original unadapted test, over all possible stage one results, also bounds its unconditional type-1 error

# Stage Two Testing by CER Method

| Intersection $H_J$ | Adapted Weights $w_j^*, j \in J$ | Weighted Pvalues $\frac{p_{j,2}^*}{w_j^*}, j \in J$ | Adjusted Cut-Off ($C_2^*$) |
|---|---|---|---|
| $H_2$ | 1 | 0.0371 | 0.0245 |
| $H_3$ | 1 | 0.0012 | 0.0245 |
| $H_4$ | 1 | 0.0195 | 0.0245 |
| $H_2 \cap H_3$ | 0.625, 0.375 | 0.0594, 0.0032 | 0.0208 |
| $H_2 \cap H_4$ | 0.75, 0.25 | 0.0495, 0.0780 | 0.0277 |
| $H_3 \cap H_4$ | 0.417, 0.583 | 0.0029, 0.0334 | 0.0296 |
| $H_2 \cap H_3 \cap H_4$ | 0.5, 0.25, 0.25 | 0.0742, 0.0048, 0.0780 | 0.0275 |

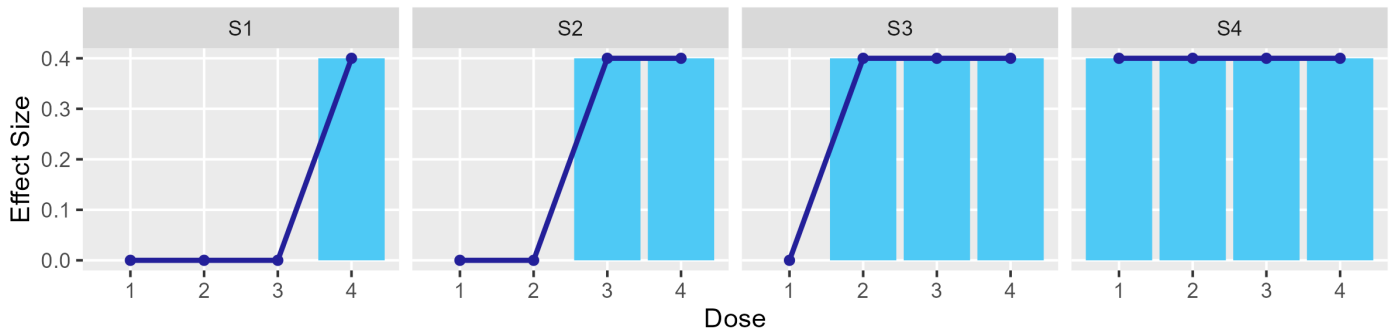
- Cumulative stage two p-values:
  $\{p_{2,2}^* = 0.0371, p_{3,2}^* = 0.0012, p_{4,2}^* = 0.0195\}$
- $H_3$ is rejected by the closed test

# Simulation Study of the Two Methods



# Scenarios and Adaptive Decision Rules

- 100 subjects/arm; 90% power at $\delta_j = 0.4, \sigma_j = 1,$ for $j = 1, \ldots 8$
- Four alternative hypothesis scenarios



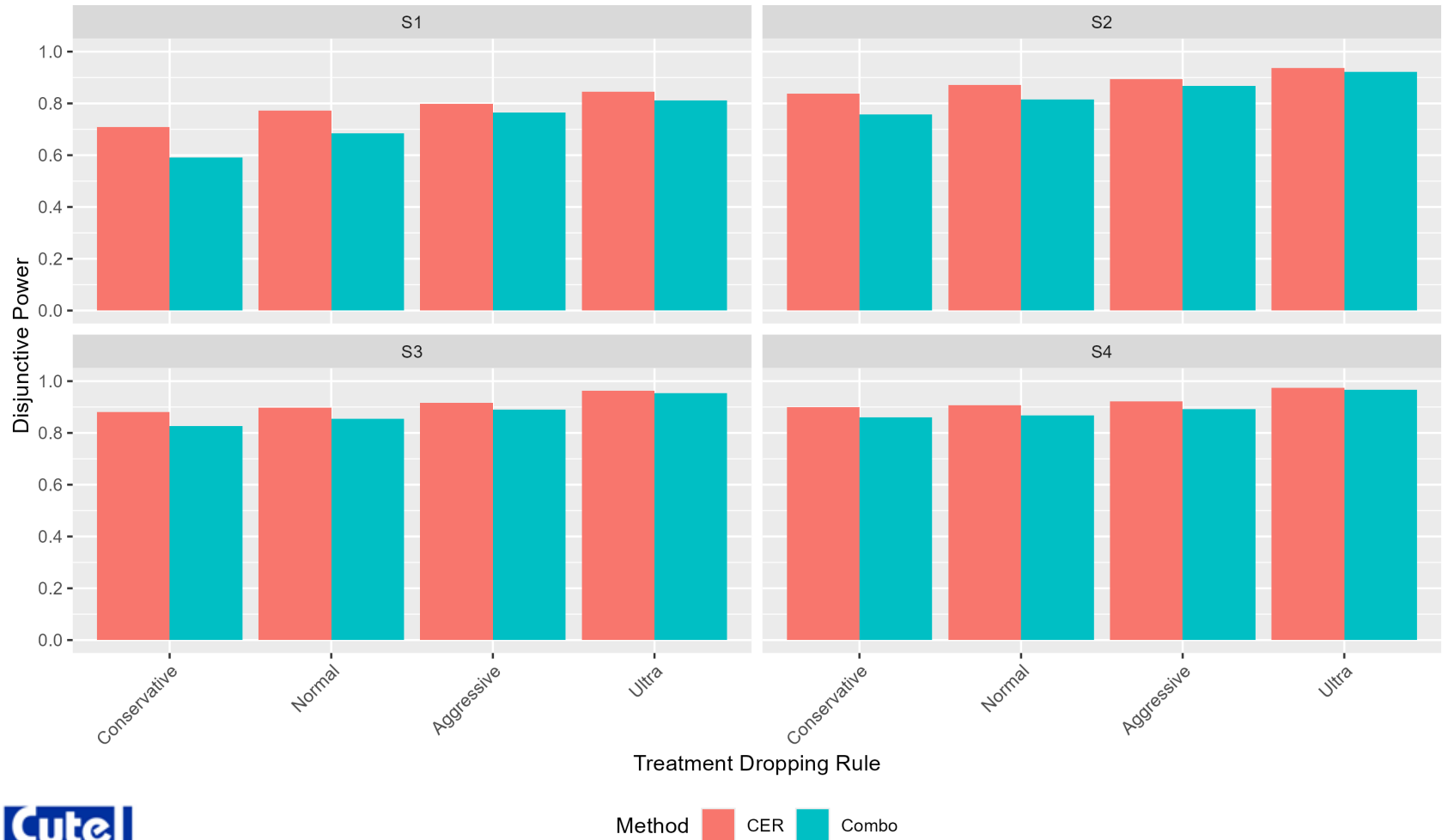

- Four decision rules for dropping non-performing doses at stage one

| | |
|---|---|
| Conservative rule | Drop doses with p-value $> 0.75$ |
| Normal rule | Drop doses with p-value $> 0.5$ |
| Aggressive rule | Drop doses with p-value $> 0.25$ |
| Ultra Aggressive rule | Select dose with smallest p-value |

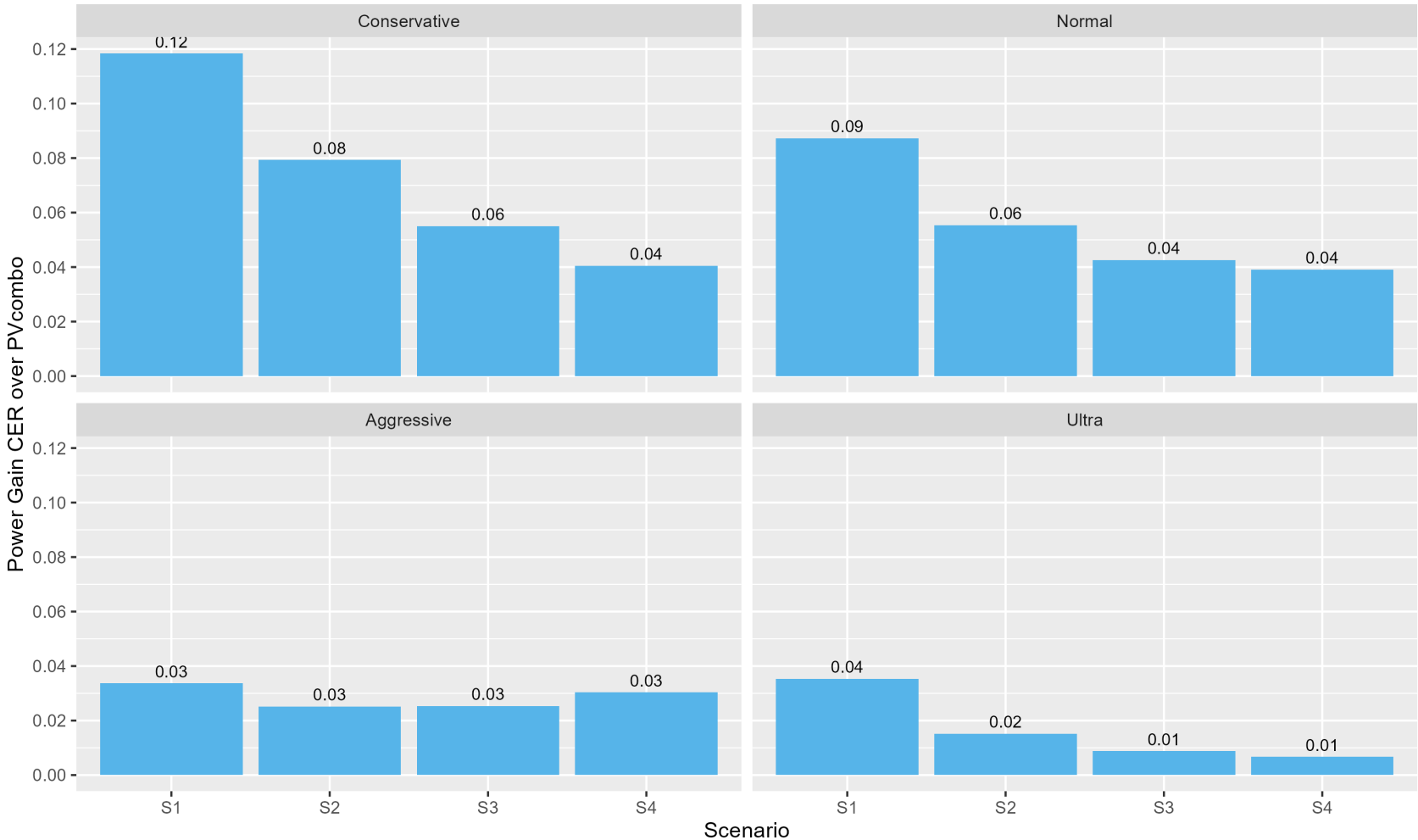
- Subjects re-assigned from dropped doses to continuing doses

# Disjunctive Power: CER vs PVcombo

Results from 10,000 Simulated Trials

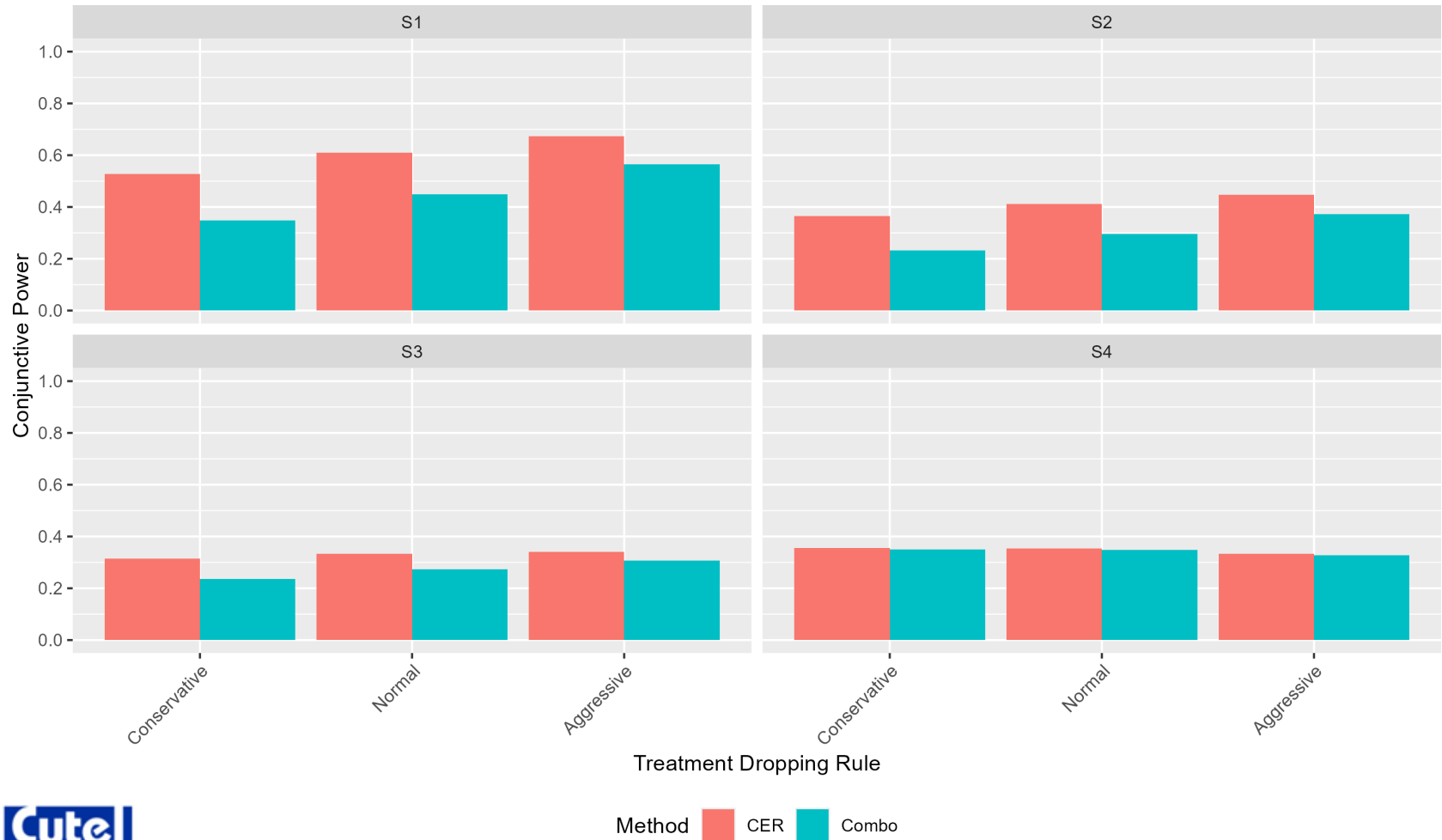


# Gain in Disjunctive Power: CER - PVcombo

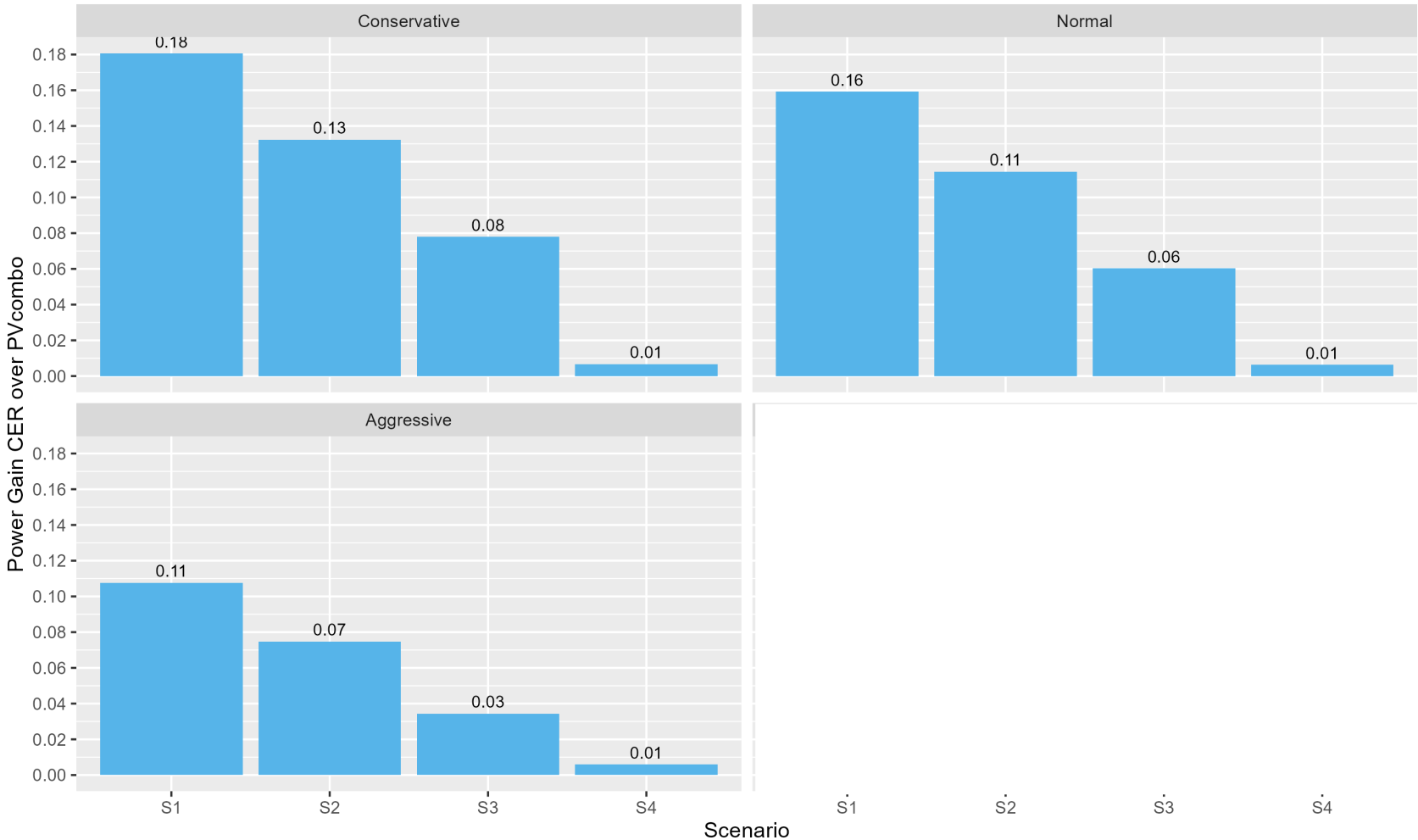


# Conjunctive Power: CER vs PVcombo

Results from 10,000 Simulated Trials

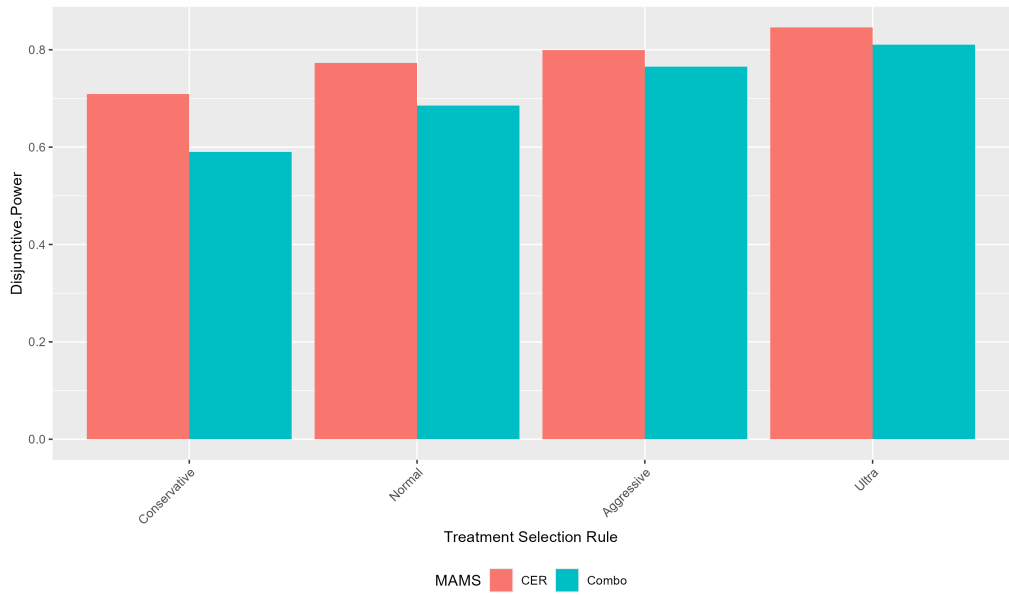


# Gain in Conjunctive Power: CER - PVcombo



# Conclusions from Simulation Results

1. Power of CER exceeds that of PVcombo everywhere
   - Up to 12% excess disjunctive and 18% excess conjunctive
2. Power gain of CER over PVcombo decreases with increasing aggressiveness for dropping doses
3. For both methods, absolute power increases with increasing aggressiveness for dropping doses



# Plausible Explanation for Greater Power of CER

## Consonance (Hommel, Bretz, Maurer, SiM 2007)

- In a consonant MTP, rejection of the global intersection hypothesis implies rejection of at least one individual hypothesis
- It follows that consonant MTPs have greater power for rejecting individual hypotheses than non-consonant ones
- In our example, PVcombo is not consonant whereas CER is consonant if doses are not dropped. This explains:
  - greater power of CER over PVcombo
  - decline in power gain with increasingly aggressive dose dropping rules
- In general CER tests lose consonance less often than PVcombo tests

# Summary

- Graphs are a natural and intuitive way to assign weights to multiple hypotheses in a MTP
- Developed initially by Bretz et al (SiM 2009) for single stage designs with Bonferroni based tests
- Extended to adaptive two-stage adaptive designs that allow early stopping, dropping of hypotheses, alteration of graphs, sample size re-assessment
- Utilize nonparametric, parametric and mixed tests depending on knowledge of correlations among the hypotheses
- Established that CER method dominates over PVcombo method over wide range of scenarios and decision rules

# Publication


RESEARCH ARTICLE | Open Access

## Graph Based, Adaptive, Multiarm, Multiple Endpoint, Two-Stage Designs

Cyrus Mehta, Ajoy Mukhopadhyay, Martin Posch

First published: 02 December 2025 | https://doi.org/10.1002/sim.70237 | VIEW METRICS

`AdaptGMCP`: R software available from
`https://github.com/Cytel-Inc/AdaptiveGMCP`