
New Developments in East: Design and Simulation of Trials with Multiple Treatment Arms

**Joint Statistical Meetings
Miami, Florida
August 3, 2011**

**Cyrus R. Mehta, Ph.D
Cytel Inc., Cambridge, MA**

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

Outline of Workshop

Multiple Comparisons Procedures in SiZ^(TM)

- Overview of SiZ^(TM)
- Brief introduction to MCP and MCP module in SiZ^(TM)
- Motivating example: Alzheimers disease
- Parametric tests (Dunnett, Step-Down Dunnett)
- Nonparametric tests (Bonferroni, Holms, Hochberg, Hommel, Fixed Sequence, Fall Back)
- Principle of closed testing
- Short-cuts to closed testing

New Architecture for East-6

- Integration of SiZ^(TM), East^(R) and modules for adaptive multi-arm trials into one unified platform

Some Sources of Multiplicity

- Multiple endpoints
- Repeated significance tests
- Multiple treatment arms
- Subgroup analysis
- Variable selection in regression models

Error Rates

- There is a ‘family’ of m inferences
- Parameters are $\delta_1, \delta_2, \dots, \delta_m$
- Null hypotheses are H_1, H_2, \dots, H_m
- Comparisonwise error rate applies to an individual hypothesis; offers ‘local control’ of type-1 error

$$\text{CER}_j = P(\text{reject } H_j | H_j \text{ is true})$$

- Familywise error rate applies to the entire family

$$\text{FWER} = P(\text{reject at least one true null hypotheses})$$

- Usually wish to control FWER at some level α

Strong and Weak Control of FWER

- Control of FWER at level α means that

$$P(\text{reject at least one true null hypotheses}) \leq \alpha \quad (1)$$

- **Strong Control** of FWER means that (1) is satisfied under all partial null hypotheses of the type $H_I = \cap_{i \in I} H_i$ for all subsets $I \subseteq \{1, 2, \dots, m\}$
- **Weak Control** of FWER means that (1) is satisfied only under some H_I , typically $I = \{1, 2, \dots, m\}$
- Strong control of FWER is a regulatory requirement if multiple statements about product efficacy are to be included in the product label

Motivation for MCP Design Software

- **Desire to run trials with more than one dose versus active comparator**
- **Sample size software for such designs is limited**
- **SiZ provide as part of a broad offering of sample size calculations. Appropriate vehicle for inclusion of MCP**

Motivating Example: Alzheimer's Disease

- Randomized, double-blind, placebo controlled, parallel group trial
- Three doses (0.3 mg, 1 mg, 2 mg) compared to placebo; daily po
- Primary endpoint: change from baseline in ADAS-cog-11 at week 24 using MMRM
- Difference from placebo expected to be between 1.5 and 2.5 units with common standard deviation $\sigma = 5$

Bonferroni Adjusted Two-Arm Trial

- Design two arm trials for 90% power; $\delta = 1.5, 2, 2.5$;
 $\sigma = 5$; 1-sided $\alpha = 0.025/3 = 0.00833$

Output Preview Area

<input checked="" type="checkbox"/>	Design	Total_Sample_Size	Power	Type1_Error	Diff_of_Means	A
<input checked="" type="checkbox"/>	Design1	602	0.9	0.00833	1.5	
<input checked="" type="checkbox"/>	Design2	338	0.9	0.00833	2	
<input checked="" type="checkbox"/>	Design3	218	0.9	0.00833	2.5	

- Doubling the sample size for a four arm trial requires between $2 \times 218 = 436$, and $2 \times 602 = 1204$ patients
- Protects type-1 error but very conservative for power because it assumes only one chance of winning, not three
- **Without MCP software, this is the only commercial option**

The MCP Module in SiZ

- Simulation based sample size calculations
- **Parametric Tests:** Single-step Dunnett and step-down Dunnett
- **Nonparametric Tests:** Bonferroni, Sidak, Weighted Bonferroni, Holm, Hochberg, Hommel, Fixed-sequence, Fall-back

Multiple Comparisons Procedures

Test: 1 Sided	Parametric	P-Value
Rejection Region: Right-Tail	<input checked="" type="checkbox"/> Dunnett's single step	<input type="checkbox"/> Bonferroni
Type - 1 Error (α): 0.025	<input type="checkbox"/> Dunnett's step-down	<input type="checkbox"/> Sidak
Number of Simulations: 10000		<input type="checkbox"/> Weighted Bonferroni
Total Sample Size (n): 320:400:20		<input type="checkbox"/> Holm's step down
		<input type="checkbox"/> Hochberg's step up
		<input type="checkbox"/> Hommel's step up
		<input type="checkbox"/> Fixed sequence
		<input type="checkbox"/> Fallback

Dunnett's Procedure

- Let $Y_{ij} \sim N(\mu_i, \sigma^2)$ be response of subject $j = 1, 2, \dots, n$ on treatment $i = 0, 1, 2, \dots, m$, where $i = 0$ denotes the control arm
- The marginal t -statistic for i th treatment effect is

$$t_i = \frac{\bar{y}_i - \bar{y}_0}{s\sqrt{2/n}}$$

- Denote the cumulative distribution function for the maximum t_i by

$$F(x|m, \nu) = P \{ \max(T_1, T_2, \dots, T_m) \leq x \}$$

- Under H_0 : $\mu_i - \mu_0 = 0$ for all i , $F(x|m, \nu)$ is multivariate- t with $\nu = (m + 1)(n - 1)$ degrees of freedom
- Compute $q_\alpha(m, \nu)$, the $(1 - \alpha)$ quantile of F , defined by

$$F(q_\alpha(m, \nu)|m, \nu) = 1 - \alpha$$

The critical value $q_\alpha(m, \nu)$ is evaluated by numerical integration

Single-Step Dunnett

- Reject every null hypothesis $H_i: \mu_i - \mu_0 = 0$ for which $t_i \geq q_\alpha(m, \nu)$
- To see why this works, let \mathcal{S} be the set of true null hypotheses. Then

$$P_{H_S} \left\{ \bigcup_{i \in \mathcal{S}} T_i \geq q_\alpha(m, \nu) \right\} \leq P_{H_0} \left\{ \bigcup_{i=1}^m T_i \geq q_\alpha(m, \nu) \right\} = 1 - F(q_\alpha(m, \nu) | m, \nu) = \alpha$$

- Equivalently, compute the **multiplicity adjusted p-values**

$$\tilde{p}_i = 1 - F(t_i | m, \nu)$$

and reject every H_i for which $\tilde{p}_i \leq \alpha$

Simulations of Single-Step Dunnett

Design | **Treatment Parameters**

Generate Means through DR Curve

Common Standard Deviation

Arm	Mean	Std.Dev.	Allocation Ratio
Control	0	5	1
1	1.5	5	1
2	2.5	5	1
3	2	5	1

Output Preview Area

<input checked="" type="checkbox"/>	Scenario ID	MCP	Global_Power	Disjunctive_Power	Conjunctive_Power	FWER	Alpha	Total_Sample_Size
<input checked="" type="checkbox"/>	Scenario1	Dunnett_Single_Step	0.8581	0.8581	0.2508	0	0.025	320
<input checked="" type="checkbox"/>	Scenario2	Dunnett_Single_Step	0.8784	0.8784	0.273	0	0.025	340
<input checked="" type="checkbox"/>	Scenario3	Dunnett_Single_Step	0.899	0.899	0.289	0	0.025	360
<input checked="" type="checkbox"/>	Scenario4	Dunnett_Single_Step	0.9118	0.9118	0.3089	0	0.025	380
<input checked="" type="checkbox"/>	Scenario5	Dunnett_Single_Step	0.9272	0.9272	0.3373	0	0.025	400

- A total sample size of 360 patients produces the desired 90% global power with Dunnett's single-step procedure
- Doubling the sample size of a Bonferroni adjusted two-arm trial would require 436-1204 patients

Step-Down Dunnett

Let $T_{(1)} \geq T_{(2)} \geq \dots \geq T_{(m)}$ denote the m order statistics

- **Step 1.** If $t_{(1)} \geq c_1 = q_\alpha(m, \nu)$, reject $H_{(1)}$ and go to the next step. Otherwise retain all hypotheses and stop.
- **Steps $i = 2, \dots, m - 1$.** If $t_{(i)} \geq c_i = q_\alpha(m - i + 1, \nu)$, reject $H_{(i)}$ and go to the next step. Otherwise retain $H_{(i)}, \dots, H_{(m)}$ and stop.
- **Step m .** If $t_{(m)} \geq c_m = q_\alpha(1, \nu)$, reject $H_{(m)}$. Otherwise retain $H_{(m)}$.

Note: Since $c_1 > c_2 > \dots > c_m$, step-down Dunnett rejects at least as many (and sometimes more) hypotheses as single step Dunnett. It controls type-1 error because it is a short-cut for a 'closed test'

Step-Down Dunnett using Adjusted P-Values

It is often more convenient to use multiplicity adjusted individual p-values, \tilde{p}_i instead of individual test statistics t_i for MCPs

- Define the significance levels $\gamma_1, \gamma_2, \dots, \gamma_m$ to be such that

$$t_{(i)} = q_{\gamma_i}(m - i + 1, \nu)$$

That is, t_i is the $(1 - \gamma_i)$ quantile of the distribution $F(\cdot | m - i + 1, \nu)$:

$$F(t_i | m - i + 1, \nu) = 1 - \gamma_i$$

- The multiplicity adjusted p-values are given by

$$\tilde{p}_{(i)} = \begin{cases} \gamma_i & \text{if } i = 1 \\ \max(\tilde{p}_{i-1}, \gamma_i) & \text{if } i = 2, \dots, m, \end{cases}$$

- Reject every hypothesis $H_{(i)}$ for which $\tilde{p}_i \leq \alpha$

Simulations of Step-Down Dunnett

Design Treatment Parameters

Multiple Comparisons Procedures

Test: 1 Sided

Rejection Region: Right-Tail

Type - 1 Error (α): 0.025

Number of Simulations: 10000

Total Sample Size (n): 360

Parametric

- Dunnett's single step
- Dunnett's step-down

P-Value

- Bonferroni
- Sidak
- Weighted Bonferroni
- Holm's step down
- Hochberg's step up
- Hommel's step up
- Fixed sequence
- Fallback

Compute Power

Create Scenario

Advanced

Clear

Output Preview Area

<input checked="" type="checkbox"/>	Scenario ID	MCP	Global_Power	Disjunctive_Power	Conjunctive_Power	FWER	Alpha	Total_Sample_Size	#_Arms	#_Simulations_Requested
<input checked="" type="checkbox"/>	Scenario1	Dunnett_Single_Step	0.8941	0.8941	0.2844	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario2	Dunnett_Step_Down	0.8995	0.8995	0.4365	0	0.025	360	4	10000

The step-down test has slightly more global and disjunctive power, and considerably more conjunctive power

Advantages and Limitations of Dunnett

Advantages

- More powerful than nonparametric procedures if assumptions are met
- Generates multiplicity adjusted confidence intervals for individual treatment effects

Limitations

- Relies on normality assumption
- Relies on homoscedasticity assumption

Dunnett's FWER under unequal variance

Generate Means through DR Curve
 Common Standard Deviation

Arm	Mean	Std.Dev.	Allocation Ratio
Control	0	5	1
1	0	5	1
2	0	5	1
3	0	10	1

Output Preview Area

<input checked="" type="checkbox"/>	Scenario ID	MCP	Global Power	Disjunctive Power	Conjunctive Power	FWER	Alpha	Total Sample Size	# Arms	# Simulations Requested
<input checked="" type="checkbox"/>	Scenario1	Dunnett_Single_Step	0.02618	0	0	0.02618	0.025	360	4	100000
<input checked="" type="checkbox"/>	Scenario2	Dunnett_Step_Down	0.02575	0	0	0.02575	0.025	360	4	100000
<input checked="" type="checkbox"/>	Scenario3	Bonferroni	0.02346	0	0	0.02346	0.025	360	4	100000
<input checked="" type="checkbox"/>	Scenario4	Sidak	0.02325	0	0	0.02325	0.025	360	4	100000

In 100,000 simulations FWER was not preserved for Dunnett but was preserved for Bonferroni and Sidak (two non-parametric MCPs)

Nonparametric MCPs

- Do not require any distributional assumptions
- Very easy to compute multiplicity adjusted p-values for the individual hypotheses
- Bonferroni and Sidak are single-step MCPs
- Holms is a step-down MCP (start with the biggest effect and work your way down)
- Hochberg and Hommel are step-up MCPs (start with the smallest effect and work your way up)
- Holms, Hochberg and Hommel MCPs are short cuts for performing corresponding closed tests
- Fixed sequence and fall back procedures test each individual hypothesis in a pre-specified order

Bonferroni and Sidak

- Let p_1, p_2, \dots, p_m be the m marginal p-values
- The Bonferroni procedure rejects any H_i for which

$$p_i \leq \frac{\alpha}{m}$$

The multiplicity adjusted p-values for the Bonferroni procedure are

$$\tilde{p}_i = \min(1, mp_i), \quad i = 1, 2, \dots, m$$

- The Sidak procedure rejects and H_i for which

$$p_i \leq 1 - (1 - \alpha)^{\frac{1}{m}}$$

The multiplicity adjusted p-values for the Sidak procedure are

$$\tilde{p}_i = 1 - (1 - p_i)^{\frac{1}{m}}, \quad i = 1, 2, \dots, m$$

Why FWER is Preserved

- Let \mathcal{S} denote the set of true null hypotheses. Then the Bonferroni result is obtained by

$$P_{\mathcal{S}} \bigcup_{i \in \mathcal{S}} (p_i \leq \frac{\alpha}{m}) \leq P_{H_0} \bigcup_{i=1}^m (p_i \leq \frac{\alpha}{m}) \leq \sum_{i=1}^m P_{H_0} (p_i \leq \frac{\alpha}{m}) = \alpha$$

- The Sidak result is obtained as follows. Let $k = 1 - (1 - \alpha)^{\frac{1}{m}}$. Then

$$P_{\mathcal{S}} \bigcup_{i \in \mathcal{S}} (p_i \leq k) \leq P_{H_0} \bigcup_{i=1}^m (p_i \leq k) = 1 - P_{H_0} \bigcap_{i=1}^m (p_i \geq k) \leq 1 - \prod_{i=1}^m (1 - k)^m = \alpha$$

The Sidak bound is sharper than the Bonferroni but requires the product assumption

Comparing the Power of Dunnett, Bonferroni and Sidak Procedures

Design Treatment Parameters

Generate Means through DR Curve

Common Standard Deviation

Arm	Mean	Std.Dev.	Allocation Ratio
Control	0	5	1
1	1.5	5	1
2	2.5	5	1
3	2	10	1

Compute Power Create Scenario | Advanced | Clear

Output Preview Area

<input checked="" type="checkbox"/>	Scenario ID	MCP	Global Power	Disjunctive Power	Conjunctive Power Δ	FWER	Alpha	Total Sample Size	#_Arms	#_Simulations Requested
<input checked="" type="checkbox"/>	Scenario1	Dunnett_Single_Step	0.7145	0.7145	0.0809	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario2	Dunnett_Step_Down	0.7122	0.7122	0.1706	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario3	Bonferroni	0.8603	0.8603	0.1185	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario4	Sidak	0.8699	0.8699	0.1147	0	0.025	360	4	10000

- Dunnett loses power because of heteroscedasticity
- Sidak is slightly more powerful than Bonferroni

The Weighted Bonferroni

- Let $w_i < 1$ be the fraction of α allocated to testing H_i , where $\sum_{i=1}^m w_i = 1$
- The weighted Bonferroni procedure rejects any H_i for which

$$p_i \leq w_i \alpha$$

- Equivalently the adjusted p-value is computed as

$$\tilde{p}_i = \min\left(1, \frac{p_i}{w_i}\right), \quad i = 1, 2, \dots, m$$

and H_i is rejected if $\tilde{p}_i \leq \alpha$

Note: The regular Bonferroni is a special case of the weighted Bonferroni in which $w_i = \frac{1}{m}$ for all i

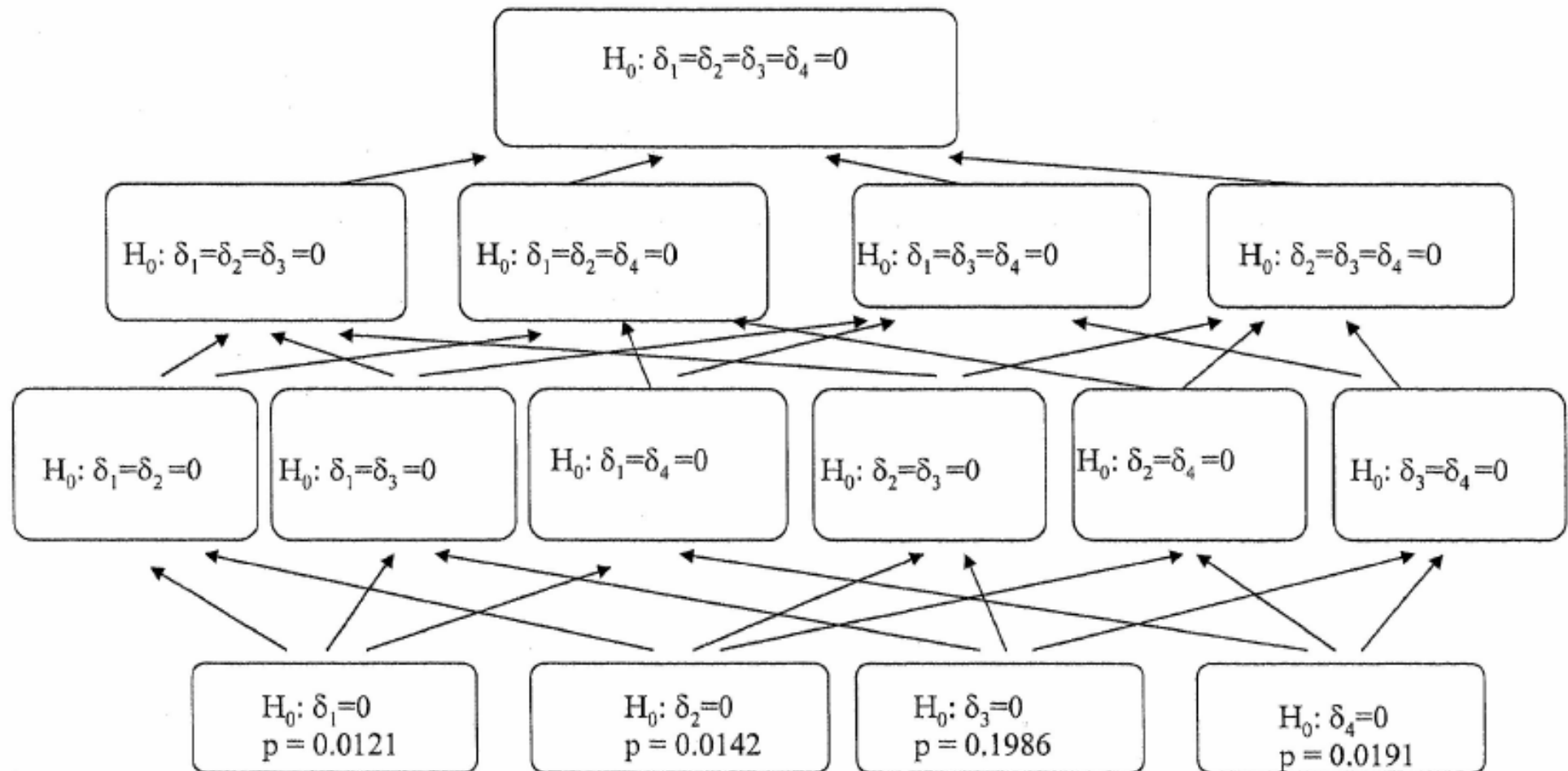
Can we improve on Bonferroni for testing individual hypotheses?

- Suppose we are testing m individual hypotheses
- Let p_1, p_2, \dots, p_i be the individual p-values
- The Bonferroni procedure rejects each H_i for which $p_i \leq \frac{\alpha}{m}$. Can we do better?
- Yes! By applying closed testing, we don't require the same strict criterion for every individual hypothesis

Closed Testing of Individual Hypotheses

- The closed testing principle is usually invoked for testing the individual members of a family of hypotheses H_1, H_2, \dots, H_m
- Suppose we wish to test each H_i while assuring strong control of FWER at level α :
 1. Construct the closed set consisting of all possible intersections of the individual hypotheses of the form $H_{i_1} \cap H_{i_2} \cap \dots \cap H_{i_q}$ for all $q = 1, 2, \dots, m$
 2. **Specify a local level- α test for each member of the closed set**
 3. An individual H_i may be rejected with strong control of FWER at level α if both these conditions hold:
 - H_i is rejected by its local level- α test
 - All intersection hypotheses that contain H_i are also rejected by their local level- α tests

Example of Closed Testing



Acknowledgement: This slide and others like it in this unit have been taken from Peter Westfall's notes

Testing the Intersection Hypotheses

The key difference between one closed testing procedure and another is the method used to test intersection hypotheses of the form $H_{i_1} \cap H_{i_2} \cap \dots \cap H_{i_q}$

- Using Dunnett's test for the intersection hypotheses leads to the step-down Dunnett procedure
- Using the Bonferroni test for the intersection hypotheses leads to the step-down Holms procedure
- Using Simes test for the intersection hypotheses leads to the step-up Hochberg and Hommel procedures

Closed Testing Using Bonferroni for the Intersection Hypotheses

- Form the closed set of all intersection hypotheses
- Reject any intersection hypothesis $\cap_{j=1}^q H_{i_j}$ if

$$\min(p_{i_1}, p_{i_2}, \dots, p_{i_q}) \leq \frac{\alpha}{q}$$

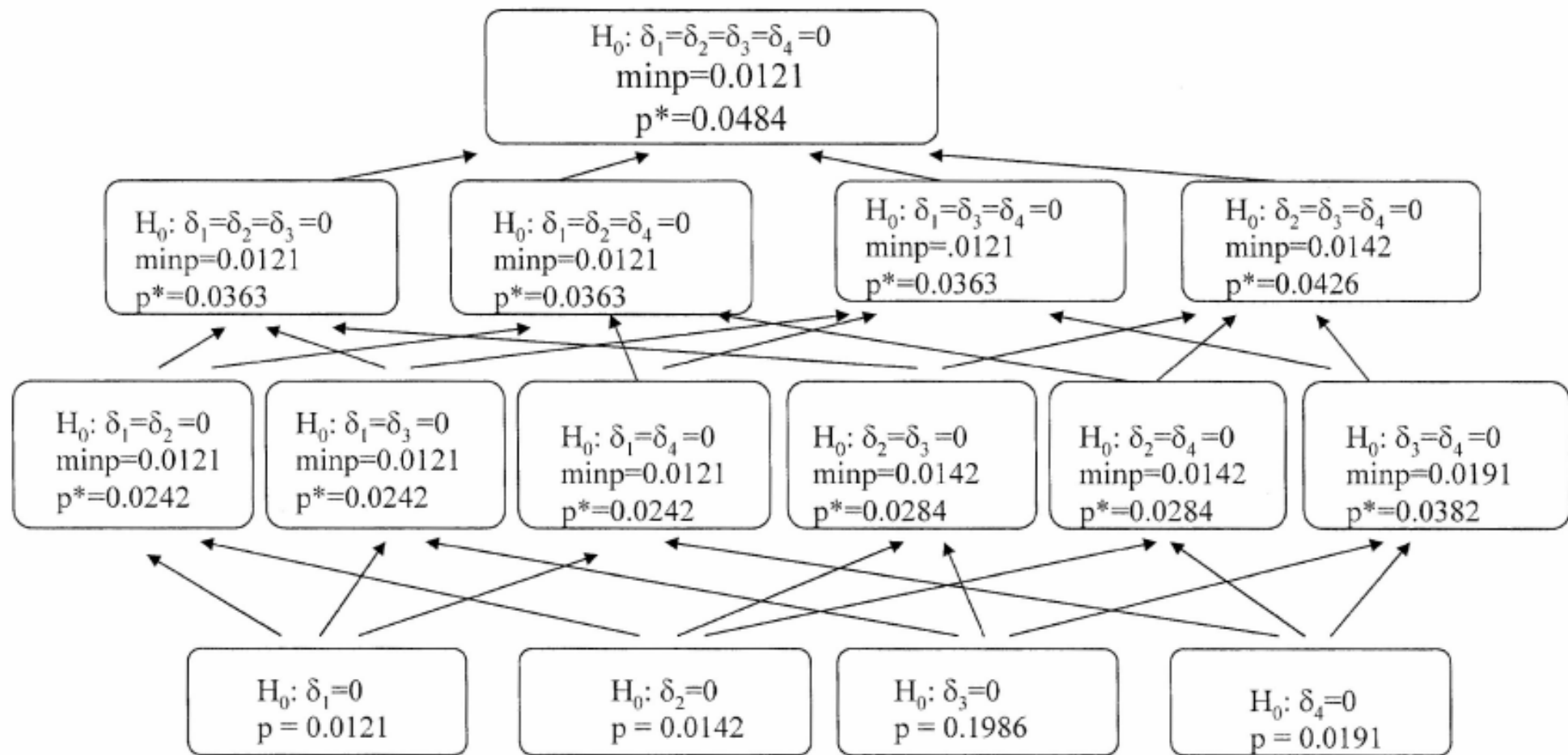
or, define the ‘adjusted p-value’

$p^* = q \times \min(p_{i_1}, p_{i_2}, \dots, p_{i_q})$ and reject $\cap_{j=1}^q H_{i_j}$ if

$$p_* \leq \alpha$$

- Better than pure Bonferroni. Every individual hypothesis need not be tested at level $\frac{\alpha}{m}$

Example of Closed Testing with Bonferroni



Only $\delta_1 = 0$ is rejected by pure Bonferroni. But $\delta_j = 0$ for $j = 1, 2, 4$ are all rejected by closed testing with Bonferroni

Bonferroni-Holm Shortcut

- Closed testing with Bonferroni is equivalent to the following short-cut:
 - Order the p-values for the individual hypotheses in ascending order as $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(m)}$
 - If $p_{(1)} \leq \alpha/m$ reject $H_{(1)}$ and continue testing; else stop
 - If $p_{(2)} \leq \alpha/(m - 1)$ reject $H_{(2)}$ and continue testing; else stop
 - In general, reject $p_{(i)}$ if $p_{(i)} \leq \alpha/(m - i + 1)$ and $H_{(1)}, H_{(2)}, \dots, H_{(i-1)}$ have already been rejected
- This is equivalent to closed testing with Bonferroni but requires at most m tests instead of searching through the entire tree
- $p_{(1)} = 0.0121 < \alpha/4 = 0.0125$; $p_{(2)} = 0.0142 < \alpha/3 = 0.0167$;
 $p_{(3)} = 0.0191 < \alpha/2 = 0.025$; but $p_{(4)} > \alpha = 0.05$

Power Comparison: Bonferroni vs Holms

Output Preview Area

<input checked="" type="checkbox"/>	Scenario ID	MCP	Global Power	Disjunctive Power	Conjunctive Power	FWER	Alpha	Total_Sample_Size	#_Arms	#_Simulations_Requested
<input checked="" type="checkbox"/>	Scenario1	Bonferroni	0.8877	0.8877	0.2721	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario2	Holm_Step_Down	0.8813	0.8813	0.4309	0	0.025	360	4	10000

Not much difference in global power. But there are significant differences in individual power

Arm	Mean	Std.Dev.	Allocation Ratio	Pure Bonferroni		Bonferroni-Holms	
				Arm	Power	Arm	Power
Control	0	5	1	Control	N.A.	Control	N.A.
1	1.5	5	1	1	0.3484	1	0.4726
2	2.5	5	1	2	0.8255	2	0.8435
3	2	5	1	3	0.609	3	0.6782

Simes Test for Intersection Hypotheses

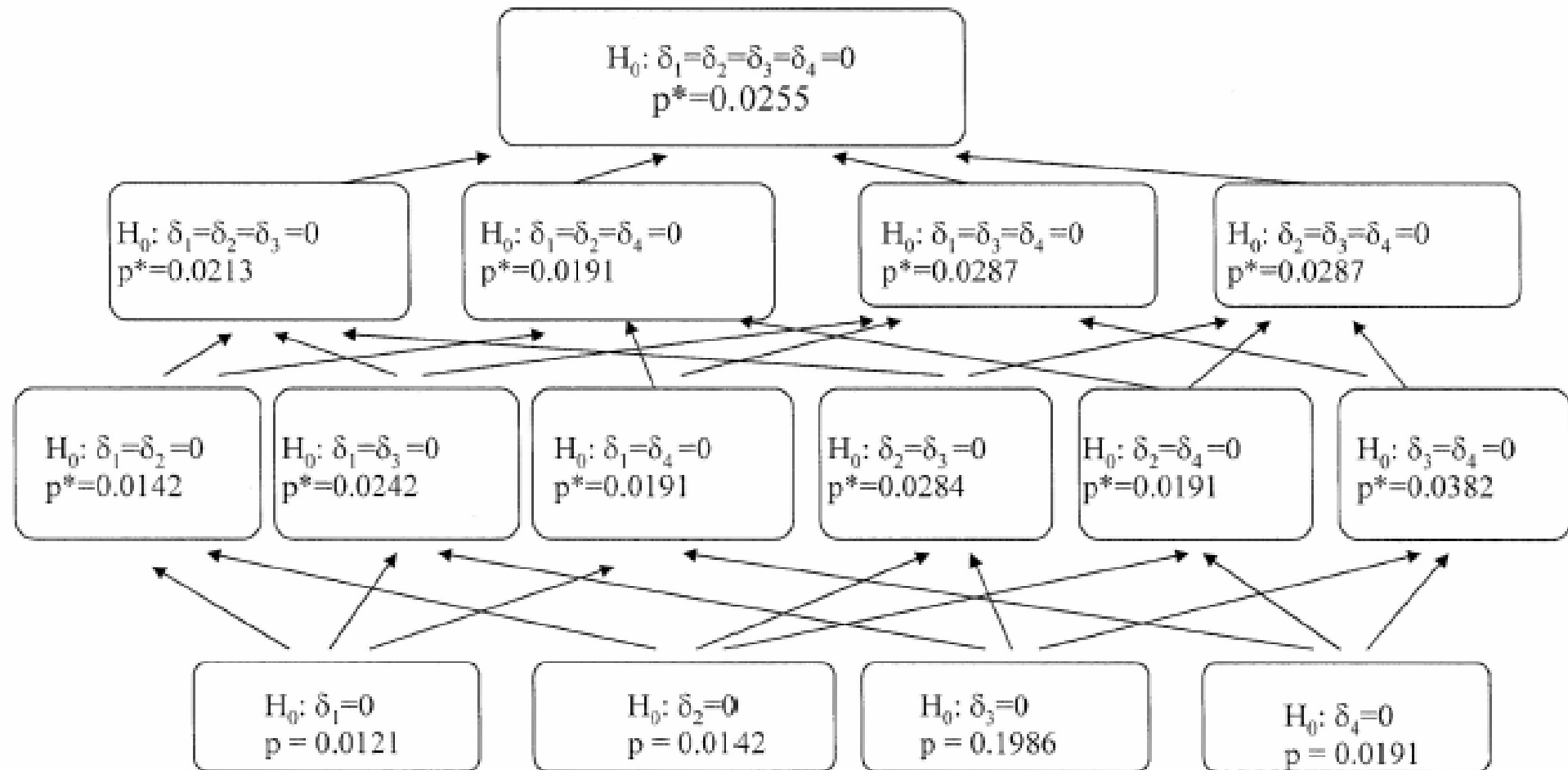
- Suppose we have k individual hypotheses H_1, H_2, \dots, H_k and wish to test the intersection hypothesis $H = \cap_{j=1}^k H_j$ at level α
- Let $p_{(1)} \leq p_{(2)}, \dots \leq p_{(k)}$ denote the k ordered p-values
- The Bonferroni test rejects H only if $p_{(1)} \leq \alpha/k$
- Simes test gives you more chances to reject H . You can reject H at level α if

$$p_{(j)} \leq \frac{j\alpha}{k}, \text{ for any } j = 1, 2, \dots, k$$

- Thus, if the smallest p-value is not small enough to get past α/k , maybe the second-smallest can get past $2\alpha/k$, and so on

In this case the adjusted p-values are $(kp^{(1)}, kp^{(2)}/2, kp^{(3)}/3 \dots p^{(k)})$.
Reject H if any of them is $< \alpha$

Example of Closed testing with Simes



This is known as Hommel's test. It is more powerful than Bonferroni-Holms

Hochberg's Shortcut to Closed Testing with Simes

- Order the p-values in ascending order as $p_{(1)} \leq p_{(2)} \cdots \leq p_{(k)}$
- If $p_{(k)} \leq \alpha$, reject all k hypotheses; otherwise retain $H_{(k)}$ and continue testing
- If $p_{(k-1)} \leq \alpha/2$, reject all $k - 1$ hypotheses; otherwise retain $H_{(k-1)}$ and continue testing
- In general, continue testing and retaining hypotheses until the first i is reached such that $p_{(i)} \leq \alpha/(k - i + 1)$. The stop further testing and reject all $H_{(i)}, H_{(i-1)}, \dots, H_{(1)}$

For the example on the previous slide, $H_{(4)} : \delta_3 = 0$ is retained since $p_{(4)} > \alpha = 0.05$. But all other hypotheses are rejected since $p_{(3)} < \alpha/2$

Hochberg's procedure is not exactly equivalent to Hommel's. It is slightly stricter and hence more conservative

Comparing Power for Bonferroni, Holms, Hochberg and Hommel Procedures

<input checked="" type="checkbox"/>	Scenario ID	MCP	Global_Power	Disjunctive_Power	Conjunctive_Power	FWER	Alpha	Total_Sample_Size	#_Arms	#_Simulations_Requested
<input checked="" type="checkbox"/>	Scenario1	Bonferroni	0.88738	0.88738	0.27587	0	0.025	360	4	100000
<input checked="" type="checkbox"/>	Scenario2	Holm_Step_Down	0.88732	0.88732	0.43673	0	0.025	360	4	100000
<input checked="" type="checkbox"/>	Scenario3	Hochberg_Step_Up	0.89228	0.89228	0.45273	0	0.025	360	4	100000
<input checked="" type="checkbox"/>	Scenario4	Hommel_Step_Up	0.89514	0.89514	0.45207	0	0.025	360	4	100000

Pure Bonferroni		Bonferroni-Holms		Hochberg		Hommel	
Arm	Power	Arm	Power	Arm	Power	Arm	Power
Control	N.A.	Control	N.A.	Control	N.A.	Control	N.A.
1	0.35058	1	0.47823	1	0.49421	1	0.49374
2	0.82962	2	0.84946	2	0.85649	2	0.85722
3	0.60603	3	0.68022	3	0.69625	3	0.69663

Small differences in global and disjunctive power. Larger differences in individual power

Fixed Sequence Testing

- Assume H_1, H_2, \dots, H_m are ordered hypotheses; i.e., $\mu_i \geq \mu_{i-1}$, $i = 1, 2, \dots, m$
- Let p_1, p_2, \dots, p_m be the associated raw p-values
 - Step 1. If $p_1 < \alpha$, reject H_1 and go to the next step. Otherwise retain all hypotheses and stop
 - Step $i = 2, \dots, m - 1$. If $p_i < \alpha$, reject H_i and go to the next step. Otherwise retaining all the remaining hypotheses and stop
 - Step m . If $p_m < \alpha$, reject H_m ; otherwise retain it.
- The adjusted p -value for H_i is given by

$$\tilde{p}_i = \max(p_1, \dots, p_i), i = 1, \dots, m.$$

- More powerful than other procedures if ordering is correct
- Closed under fixed sequence testing of each intersection hypotheses

Example 1: Testing sequence is correct

Design | **Treatment Parameters**

Generate Means through DR Curve

Common Standard Deviation

Arm	Mean	Std.Dev.	Allocation Ratio	Test Sequence
Control	0	5	1	
1	1.5	5	1	3
2	2.5	5	1	1
3	2	5	1	2

Compute Power

Create Scenario

Advanced

Clear

Output Preview Area

<input checked="" type="checkbox"/>	Scenario ID	MCP	Global Power	Disjunctive Power	Conjunctive Power	FWER	Alpha	Total Sample Size	#_Arms	#_Simulations Requested
<input checked="" type="checkbox"/>	Scenario1_1	Dunnett_Step_Down	0.8958	0.8958	0.4444	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_2	Bonferroni	0.8903	0.8903	0.2693	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_3	Holm_Step_Down	0.8905	0.8905	0.4408	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_4	Hochberg_Step_Up	0.8901	0.8901	0.4471	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_5	Hommel_Step_Up	0.8951	0.8951	0.4526	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_6	Fixed_Sequence	0.9177	0.9177	0.4548	0	0.025	360	4	10000

If the order of testing is correctly specified, the fixed sequence procedure is the most powerful

Example 2: Testing sequence is incorrect

Design | **Treatment Parameters**

Generate Means through DR Curve

Common Standard Deviation

Arm	Mean	Std.Dev.	Allocation Ratio	Test Sequence
Control	0	5	1	
1	1.5	5	1	1
2	2.5	5	1	2
3	2	5	1	3

Compute Power

Create Scenario

Advanced

Clear

Output Preview Area

<input checked="" type="checkbox"/>	Scenario ID	MCP	Global Power	Disjunctive Power	Conjunctive Power	FWER	Alpha	Total_Sample_Size	#_Arms	#_Simulations_Requested
<input checked="" type="checkbox"/>	Scenario1_1	Dunnett_Step_Down	0.8973	0.8973	0.4374	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_2	Bonferroni	0.8823	0.8823	0.2704	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_3	Holm_Step_Down	0.8939	0.8939	0.4433	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_4	Hochberg_Step_Up	0.8914	0.8914	0.4404	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_5	Hommel_Step_Up	0.8932	0.8932	0.4568	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_6	Fixed_Sequence	0.515	0.515	0.4454	0	0.025	360	4	10000

Fixed sequence procedure loses considerable power if testing sequence is incorrect

The Fall Back Procedure

- Assume H_1, H_2, \dots, H_m are ordered hypotheses, w_1, w_2, \dots, w_m are pre-specified weights with $\sum_{i=1}^m w_i = 1$, and p_1, p_2, \dots, p_m are the associated raw p-values. The testing proceeds as follows:
 - Step 1. Test H_1 at $\alpha_1 = w_1\alpha$. If $p_1 \leq \alpha_1$, reject H_1 ; otherwise retain it and go to the next step.
 - Step $i = 2, \dots, m$. Test H_i at $\alpha_i = \alpha_{i-1} + w_i\alpha$ if H_{i-1} is rejected and at $\alpha_i = w_i\alpha$ if H_{i-1} is retained. If $p_i \leq \alpha_i$, reject H_i ; otherwise retain it and go to the next step.
- Provides option to continue even if a hypothesis is retained. Hence good insurance policy in case incorrect testing order was specified
- Specializes to fixed sequence test if $w_1 = 1$ and $w_2 = \dots = w_m = 0$
- Was shown by Wiens and Dmitreinko (2005) to be a closed test

Example 3: Performance of Fall Back Test under Incorrect Testing Sequence

Arm	Mean	Std.Dev.	Allocation Ratio	Proportion of Alpha	Test Sequence
Control	0	5	1		
1	1.5	5	1	0.333	1
2	2.5	5	1	0.333	2
3	2	5	1	0.333	3

Test H_1 at level $\alpha_1 = \alpha/3$. If H_1 is rejected, test H_2 at level $\alpha_2 = \alpha_1 + \alpha/3$, otherwise test H_2 at level $\alpha_2 = \alpha/3$. If H_2 is rejected test H_3 at level $\alpha_2 + \alpha/3$, otherwise test H_3 at level $\alpha/3$

Output Preview Area										
<input checked="" type="checkbox"/>	Scenario ID	MCP	Global Power	Disjunctive Power	Conjunctive Power	FWER	Alpha	Total_Sample_Size	#_Arms	#_Simulations_Requested
<input checked="" type="checkbox"/>	Scenario1_1	Hommel_Step_Up	0.899	0.899	0.4569	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_2	Fixed_Sequence	0.512	0.512	0.4463	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_3	Fallback	0.8888	0.8888	0.3145	0	0.025	360	4	10000

Fall back procedure is almost as good as Hommel despite guessing treatment order incorrectly

Example 4: Performance of Fall Back Test under Correct Testing Sequence

Arm	Mean	Std.Dev.	Allocation Ratio	Proportion of Alpha	Test Sequence
Control	0	5	1		
1	1.5	5	1	0.333	3
2	2.5	5	1	0.333	1
3	2	5	1	0.333	2

Output Preview Area

<input checked="" type="checkbox"/>	Scenario ID	MCP	Global Power	Disjunctive Power	Conjunctive Power	FWER	Alpha	Total Sample Size	#_Arms	#_Simulations Requested
<input checked="" type="checkbox"/>	Scenario1_1	Hommel_Step_Up	0.8956	0.8956	0.456	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_2	Fixed_Sequence	0.9178	0.9178	0.452	0	0.025	360	4	10000
<input checked="" type="checkbox"/>	Scenario1_3	Fallback	0.8925	0.8925	0.4118	0	0.025	360	4	10000

In this example the fall back test is almost as good as the fixed sequence test when the ordering is correctly specified and is superior to the fixed sequence test when the ordering is incorrectly specified