A graphical approach to sequentially rejective multiple test procedures

Martin Posch
Center for Medical Statistics, Informatics and Intelligent Systems
Medical University of Vienna

Joint work with Frank Bretz, Willi Maurer, Werner Brannath

μToss, Berlin 2010
Sequentially rejective, weighted Bonferroni type tests

- Applied in clinical trials with multiple treatment arms, subgroups and endpoints
- Bonferroni-Holm Test, Fixed Sequence Test, Fallback Test, Gatekeeping Tests, ...
- Allow to map the difference in importance as well as the relationship between research questions onto the multiple test procedure.
- However: The testing procedure can be technical and often hard to communicate.

Parallell gatekeeping: Testing $\mathcal{F}_1 = \{H_1, H_2\}$, $\mathcal{F}_2 = \{H_3, H_4\}$

Rejection of hypotheses in the family $\mathcal{F}_2 = \{H_3, H_4\}$ is only of interest if at least one of the hypotheses in the family $\mathcal{F}_1 = \{H_1, H_2\}$ can be rejected
Table II. Decision matrix for the parallel Bonferroni gatekeeping procedure.

<table>
<thead>
<tr>
<th>Intersection hypothesis</th>
<th>( P )-values for intersection hypotheses</th>
<th>( H_1 )</th>
<th>( H_2 )</th>
<th>( H_3 )</th>
<th>( H_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_{1111} )</td>
<td>( p_{1111} = \min(p_1/0.9, p_2/0.1) )</td>
<td>( p_{1111} )</td>
<td>( p_{1111} )</td>
<td>( p_{1111} )</td>
<td>( p_{1111} )</td>
</tr>
<tr>
<td>( H_{1110} )</td>
<td>( p_{1110} = \min(p_1/0.9, p_2/0.1) )</td>
<td>( p_{1110} )</td>
<td>( p_{1110} )</td>
<td>( p_{1110} )</td>
<td>( 0 )</td>
</tr>
<tr>
<td>( H_{1101} )</td>
<td>( p_{1101} = \min(p_1/0.9, p_2/0.1) )</td>
<td>( p_{1101} )</td>
<td>( p_{1101} )</td>
<td>( 0 )</td>
<td>( p_{1101} )</td>
</tr>
<tr>
<td>( H_{1100} )</td>
<td>( p_{1100} = \min(p_1/0.9, p_2/0.1) )</td>
<td>( p_{1100} )</td>
<td>( p_{1100} )</td>
<td>( 0 )</td>
<td>( 0 )</td>
</tr>
<tr>
<td>( H_{1011} )</td>
<td>( p_{1011} = \min(p_1/0.9, p_3/0.05, p_4/0.05) )</td>
<td>( p_{1011} )</td>
<td>( 0 )</td>
<td>( p_{1011} )</td>
<td>( p_{1011} )</td>
</tr>
<tr>
<td>( H_{1010} )</td>
<td>( p_{1010} = \min(p_1/0.9, p_3/0.1) )</td>
<td>( p_{1010} )</td>
<td>( 0 )</td>
<td>( p_{1010} )</td>
<td>( 0 )</td>
</tr>
<tr>
<td>( H_{1001} )</td>
<td>( p_{1001} = \min(p_1/0.9, p_4/0.1) )</td>
<td>( p_{1001} )</td>
<td>( 0 )</td>
<td>( 0 )</td>
<td>( p_{1001} )</td>
</tr>
<tr>
<td>( H_{1000} )</td>
<td>( p_{1000} = p_1 )</td>
<td>( p_{1000} )</td>
<td>( 0 )</td>
<td>( 0 )</td>
<td>( 0 )</td>
</tr>
<tr>
<td>( H_{0111} )</td>
<td>( p_{0111} = \min(p_2/0.1, p_3/0.45, p_4/0.45) )</td>
<td>( 0 )</td>
<td>( p_{0111} )</td>
<td>( p_{0111} )</td>
<td>( p_{0111} )</td>
</tr>
<tr>
<td>( H_{0110} )</td>
<td>( p_{0110} = \min(p_2/0.1, p_3/0.9) )</td>
<td>( 0 )</td>
<td>( p_{0110} )</td>
<td>( p_{0110} )</td>
<td>( 0 )</td>
</tr>
<tr>
<td>( H_{0101} )</td>
<td>( p_{0101} = \min(p_2/0.1, p_4/0.9) )</td>
<td>( 0 )</td>
<td>( p_{0101} )</td>
<td>( 0 )</td>
<td>( p_{0101} )</td>
</tr>
<tr>
<td>( H_{0100} )</td>
<td>( p_{0100} = p_2 )</td>
<td>( 0 )</td>
<td>( p_{0100} )</td>
<td>( 0 )</td>
<td>( 0 )</td>
</tr>
<tr>
<td>( H_{0011} )</td>
<td>( p_{0011} = \min(p_3/0.5, p_4/0.5) )</td>
<td>( 0 )</td>
<td>( 0 )</td>
<td>( p_{0011} )</td>
<td>( p_{0011} )</td>
</tr>
<tr>
<td>( H_{0010} )</td>
<td>( p_{0010} = p_3 )</td>
<td>( 0 )</td>
<td>( 0 )</td>
<td>( p_{0010} )</td>
<td>( 0 )</td>
</tr>
<tr>
<td>( H_{0001} )</td>
<td>( p_{0001} = p_4 )</td>
<td>( 0 )</td>
<td>( 0 )</td>
<td>( 0 )</td>
<td>( p_{0001} )</td>
</tr>
</tbody>
</table>

Note: The table shows \( p \)-values associated with the intersection hypotheses. The adjusted \( p \)-values for the original hypotheses \( H_1, H_2, H_3 \) and \( H_4 \) are defined as the largest \( p \)-value in the corresponding column in the right-hand panel of the table (see equation (1)).
Heuristics

**Notation**

- $H_1, \ldots, H_m$: $m$ null hypotheses.
- $p_1, \ldots, p_m$: $m$ elementary p-values
- $\alpha = (\alpha_1, \ldots, \alpha_m)$: initial allocation of the type I error rate
  \[ \alpha = \sum_{i=1}^{m} \alpha_i. \]

**“$\alpha$ Reshuffling”**

1. If a hypothesis $H_i$ can be rejected at level $\alpha_i$, reallocate its level to one of the other hypotheses (according to a prefixed rule)
2. Repeat the testing with the resulting $\alpha$ levels.
3. Go to step 1 until no hypothesis can be rejected anymore.

Does this lead to a FWE-controlling test?
Example: Bonferroni-Holm Test

\[ H_1 \xrightarrow{\alpha/2} 1 \xrightarrow{1} H_2 \xrightarrow{\alpha/2} H_1 \]
Example: Bonferroni-Holm Test ($\alpha = 0.025$)

- $H_1$ with $p_1 = 0.04$
- $H_2$ with $p_2 = 0.01$
Example: Bonferroni-Holm Test ($\alpha = 0.025$)

$H_1$  

$p_1 = 0.04$

$\frac{\alpha}{2}$  

$H_2$

$p_2 = 0.01$

$\frac{\alpha}{2}$

$1$  

$1$
Example: Bonferroni-Holm Test ($\alpha = 0.025$)

$H_1$  \[\alpha\]  $H_2$

$p_1 = 0.04$
Example: Bonferroni-Holm Test ($\alpha = 0.025$)

$H_1$

$\alpha$

$p_1 = 0.04$
Example: Parallel Gatekeeping

To the procedure of Dmitrienko et al. (2003)
Example: Parallel Gatekeeping ($\alpha = 0.025$)

- $p_1 = 0.01$
- $p_3 = 0.001$
- $p_2 = 0.005$
- $p_4 = 0.04$
Example: Parallel Gatekeeping ($\alpha = 0.025$)

$\alpha/2 \quad 1/2 \quad 1/2 \quad \alpha/2$

$p_1 = 0.01 \quad H_1 \quad 1/2 \quad 1 \quad 1/2 \quad H_2 \quad p_2 = 0.005$

$p_3 = 0.001 \quad H_3 \quad 0 \quad 0 \quad 0 \quad H_4 \quad p_4 = 0.04$
Example: Parallel Gatekeeping ($\alpha = 0.025$)

$p_1 = 0.01$

$p_3 = 0.001$

$p_2 = 0.005$

$p_4 = 0.04$
Example: Parallel Gatekeeping ($\alpha = 0.025$)

\[ p_2 = 0.005 \quad \frac{\alpha}{2} \]

\[ p_4 = 0.04 \quad \frac{\alpha}{4} \]

\[ p_3 = 0.001 \quad \frac{\alpha}{4} \]

\[ H_1 \quad \frac{1}{2} \quad \frac{1}{2} \quad \frac{1}{2} \quad \frac{1}{2} \]

\[ H_2 \]

\[ H_3 \]

\[ H_4 \]

\[ H_1 \rightarrow H_2 \]

\[ H_2 \rightarrow H_4 \]

\[ H_3 \rightarrow H_4 \]

\[ H_1 \leftarrow H_3 \]

\[ H_2 \leftarrow H_4 \]
Example: Parallel Gatekeeping ($\alpha = 0.025$)

\[
\begin{align*}
\alpha &= 0.025 \\
H_2 \ p_2 &= 0.005 \\
H_3 \ p_3 &= 0.001 \\
H_4 \ p_4 &= 0.04
\end{align*}
\]
Example: Parallel Gatekeeping ($\alpha = 0.025$)

- $H_3$: $p_3 = 0.001$
- $H_2$: $p_2 = 0.005$
- $H_4$: $p_4 = 0.04$

Graph:
- $H_3$ to $H_2$: $\frac{\alpha}{4}$
- $H_2$ to $H_4$: $\frac{1}{2}$
- $H_3$ to $H_4$: $1$

Math:
- $H_2 p_2 = 0.005$
- $H_3 p_3 = 0.001$
- $H_4 p_4 = 0.04$
Example: Parallel Gatekeeping ($\alpha = 0.025$)

$$H_2 p_2 = 0.005$$

$$H_3$$

$$H_4 p_4 = 0.04$$
Example: Parallel Gatekeeping ($\alpha = 0.025$)

$H_2$  \hspace{1cm} $p_2 = 0.005$

$H_4$  \hspace{1cm} $p_4 = 0.04$
Example: Parallel Gatekeeping ($\alpha = 0.025$)

\[ \frac{\alpha}{2} \]

\[ H_2 \]

\[ p_2 = 0.005 \]

\[ 1 \]

\[ H_4 \]

\[ \frac{\alpha}{2} \]

\[ p_4 = 0.04 \]
Example: Parallel Gatekeeping ($\alpha = 0.025$)

\[ H_4 \quad p_4 = 0.04 \]
General Definition of the Multiple Test Procedure

General definition of the multiple test

- $\alpha = (\alpha_1, \ldots, \alpha_m)$, $\sum_{i=1}^{m} \alpha_i = \alpha$, initial levels
- $G = (g_{ij}) : m \times m$ transition matrix
  $g_{ij}$ with $0 \leq g_{ij} \leq 1$, $g_{ii} = 0$ and $\sum_{j=1}^{m} g_{ij} \leq 1$ for all $i = 1, \ldots, m$.

- $g_{ij}$...fraction of the level of $H_i$ that is allocated to $H_j$.
- $G$ and $\alpha$ determine the graph and the multiple test.
The Testing Procedure

Set $J = \{1, \ldots, m\}$.

1. Select a $j$ such that $p_j \leq \alpha_j$.
   If no such $j$ exists, stop, otherwise reject $H_j$.

2. Update the graph:

   $J \rightarrow J/\{j\}$

   $\alpha_\ell \rightarrow \begin{cases} 
   \alpha_\ell + \alpha_j g_{\ell j}, & \ell \in J \\
   0, & \text{otherwise}
   \end{cases}$

   $g_{\ell k} \rightarrow \begin{cases} 
   \frac{g_{\ell k} + g_{ejg_{jk}}}{1 - g_{ejg_{je}}}, & \ell, k \in J, \ell \neq k \\
   0, & \text{otherwise}
   \end{cases}$

3. Go to step 1.
The Testing Procedure

Set $J = \{1, \ldots, m\}$.

1. Select a $j$ such that $p_j \leq \alpha_j$.
   If no such $j$ exists, stop, otherwise reject $H_j$.

2. Update the graph:

   $J \rightarrow J/\{j\}$

   $\alpha_\ell \rightarrow \left\{ \begin{array}{ll}
   \alpha_\ell + \alpha_j g_{j\ell}, & \ell \in J \\
   0, & \text{otherwise}
   \end{array} \right.$

   $g_{\ell k} \rightarrow \left\{ \begin{array}{ll}
   \frac{g_{\ell k} + g_{j\ell} g_{jk}}{1 - g_{j\ell} g_{je}}, & \ell, k \in J, \ell \neq k \\
   0, & \text{otherwise}
   \end{array} \right.$

3. Go to step 1.
The Testing Procedure

Set $J = \{1, \ldots, m\}$.

1. Select a $j$ such that $p_j \leq \alpha_j$. If no such $j$ exists, stop, otherwise reject $H_j$.

2. Update the graph:

   $$J \rightarrow J/\{j\}$$

   $$\alpha_\ell \rightarrow \begin{cases} 
   \alpha_\ell + \alpha_j g_{j\ell}, & \ell \in J \\
   0, & \text{otherwise}
   \end{cases}$$

   $$g_{\ell k} \rightarrow \begin{cases} 
   \frac{g_{\ell k} + g_{j\ell} g_{j k}}{1 - g_{j\ell} g_{j e}}, & \ell, k \in J, \ell \neq k \\
   0, & \text{otherwise}
   \end{cases}$$

3. Go to step 1.
Updating the Graph

Graph:
- Nodes: $H_1$, $H_2$, $H_3$
- Edges:
  - $g_{12}$ from $H_1$ to $H_2$
  - $g_{23}$ from $H_2$ to $H_3$
  - $g_{13}$ from $H_1$ to $H_3$
Updating the Graph

\[ g_{12} \quad g_{23} \]

\[ H_1 \quad H_2 \quad H_3 \]

\[ g_{13} + g_{12}g_{23} \]
Updating the Graph
Updating the Graph

$H_1 \rightarrow H_2 \rightarrow H_3$

$g_{12} \rightarrow H_2$

$g_{21} \rightarrow H_1$

$g_{23} \rightarrow H_3$

$g_{12}g_{21}$

$g_{13} + g_{12}g_{23}$
Updating the Graph

\[
\begin{align*}
g_{12} + g_{12}g_{23} &\quad \frac{g_{13} + g_{12}g_{23}}{1 - g_{12}g_{21}}
\end{align*}
\]
### Theorem

The initial levels $\alpha$, the transition matrix $G$ and the algorithm define a unique multiple testing procedure controlling strongly the FWER at level $\alpha$.

**Proof:**

- The graph and algorithm define weighted Bonferroni tests for all intersection hypotheses.
- The algorithm is a short cut for the resulting closed test.
Closed Testing with Weighted Bonferroni Tests

Closed Testing Procedure:

1. Define level $\alpha$ tests for all intersection hypotheses
   $$H_J = \bigcap_{i \in J} H_i, J \subseteq \{1, \ldots, m\}.$$
2. Reject $H_j$, at multiple level $\alpha$, if for all $J \subseteq \{1, \ldots, m\}$ that contain $j$ the intersection hypotheses $H_J$ can be rejected at level $\alpha$.

Weighted Bonferroni Test.

1. For each $J \subseteq \{1, \ldots, m\}$ define $\alpha^J_j$ such that $\sum_{j \in J} \alpha^J_j = \alpha$.
2. Reject $H_J$, if $p_j \leq \alpha^J_j$ for some $j \in J$. 

Fixed Sequence Test

\[ \alpha = (\alpha, 0, 0), \quad G = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \]
\[ \alpha = (\alpha_1, \alpha_2, \alpha_3), \quad \mathbf{G} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \]
Improved Fallback Procedure  (Wiens & Dmitrienko, 2005)

\[ \alpha = (\alpha_1, \alpha_2, \alpha_3), \quad G = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix} \]

\[
\begin{array}{c}
\alpha_1 \\
H_1 \\
\end{array} \quad \begin{array}{c}
1 \\
\rightarrow \\
H_2 \\
\end{array} \quad \begin{array}{c}
\alpha_2 \\
1 \\
\rightarrow \\
H_3 \\
\end{array} \quad \begin{array}{c}
\alpha_3 \\
\end{array}
\]
Yet another improved Fallback Procedure

Let $\epsilon \to 0$, see explanation below.
Shifting level between families of hypotheses (1)

Test strategy

- $H_1, H_2$ tested with Bonferroni-Holm
- $H_3$ tested (at level $\alpha$) only if $H_1$ and $H_2$ are rejected
Shifting level between families of hypotheses (2)

\[ \alpha = \left( \frac{\alpha}{2}, \frac{\alpha}{2}, 0 \right), \quad G = \begin{pmatrix}
0 & 1 - \epsilon & \epsilon \\
1 - \epsilon & 0 & \epsilon \\
0 & 0 & 0
\end{pmatrix} \]

Let \( \epsilon \to 0 \).
Let \( \epsilon \to 0 \).
Shifting level between families of hypotheses (2)

Let $\epsilon \to 0$. 
$\alpha = (\alpha, 0, 0), \quad G = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Let $\epsilon \to 0$. 
Parallel Gatekeeping (Dmitrienko, Offen & Westfall, 2003)

\[ \alpha = \left( \frac{\alpha}{2}, \frac{\alpha}{2}, 0, 0 \right), \quad \mathbf{G} = \begin{pmatrix}
0 & 0 & 0.5 & 0.5 \\
0 & 0 & 0.5 & 0.5 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
\end{pmatrix} \]
Improved Parallel Gatekeeping (Hommel, Bretz & Maurer, 2007)

\[ \alpha = \left( \frac{\alpha}{2}, \frac{\alpha}{2}, 0, 0 \right), \quad G = \begin{pmatrix}
0 & 0 & 0.5 & 0.5 \\
0 & 0 & 0.5 & 0.5 \\
\epsilon & 0 & 0 & 1 - \epsilon \\
0 & \epsilon & 1 - \epsilon & 0
\end{pmatrix} \]
When is a graph complete? 
... and cannot be improved by adding additional edges?

A sufficient condition for completeness:

- the weights of outgoing edges sum to one at each node and
- every node is accessible from any of the other nodes

If $\alpha_i > 0, i = 1, \ldots, m$, this is also a necessary condition for completeness.
How general is the procedure?

Can all consonant closed test procedures using weighted Bonferroni Tests for the intersection hypotheses be constructed with the graphical procedure?

No:

- For the general procedure we can choose weights for $2^{m-1}$ intersection hypotheses.
- The graphical procedure is defined by $m^2 + m$ parameters.
Extensions

- Multiplicity adjusted confidence bounds (Guilbaud (2008) and Strassburger and Bretz (2008))
- Adjusted p-values
Construction of Adjusted Confidence Bounds

Assumptions:

- Test for $H_i : \theta_i \leq 0$ v.s. $H_i' : \theta_i > 0$
- Let $p_i(\mu)$ denote a p-value for $H_i(\mu) : \theta_i \leq \mu$.
- $p_i(\mu)$ is increasing in $\mu$.
- $b_i(\gamma) = \inf\{\mu | p_i(\mu) > \gamma\}$ (local level $\gamma$ confidence bound)
- $I \subseteq \{1, \ldots, m\}$ ... index set of rejected hypotheses $H_i$.

The adjusted bounds

- If $I = \{1, \ldots, m\}$: $b_i^{adj} = \max\{0, b_i(\alpha_i)\}$.
- Otherwise: $b_i^{adj} = \begin{cases} 0 & \text{if } i \in I \\ b_i(\alpha'_i) & \text{otherwise.} \end{cases}$

$\alpha'_i$ ... level of hypothesis $H_i$ in the final graph.
Construction of Adjusted Confidence Bounds

Assumptions:

- Test for $H_i : \theta_i \leq 0$ v.s. $H'_i : \theta_i > 0$
- Let $p_i(\mu)$ denote a p-value for $H_i(\mu) : \theta_i \leq \mu$.
  - $p_i(\mu)$ is increasing in $\mu$.
  - $b_i(\gamma) = \inf\{\mu | p_i(\mu) > \gamma\}$ (local level $\gamma$ confidence bound)
- $I \subseteq \{1, \ldots, m\}$ ... index set of rejected hypotheses $H_i$.

The adjusted bounds

- If $I = \{1, \ldots, m\}$: $b_{i}^{adj} = \max\{0, b_i(\alpha_i)\}$.
- Otherwise: $b_{i}^{adj} = \begin{cases} 0 & \text{if } i \in I \\ b_i(\alpha'_i) & \text{otherwise.} \end{cases}$

$\alpha'_i$ ... level of hypothesis $H_i$ in the final graph.
Assumptions:

- Test for $H_i : \theta_i \leq 0$ v.s. $H'_i : \theta_i > 0$
- Let $p_i(\mu)$ denote a p-value for $H_i(\mu) : \theta_i \leq \mu$.
- $p_i(\mu)$ is increasing in $\mu$.
- $b_i(\gamma) = \inf\{\mu | p_i(\mu) > \gamma\}$ (local level $\gamma$ confidence bound)
- $I \subseteq \{1, \ldots, m\}$ ... index set of rejected hypotheses $H_i$.

The adjusted bounds

- If $I = \{1, \ldots, m\}$: $b_{i,\text{adj}} = \max\{0, b_i(\alpha_i)\}$.
- Otherwise: $b_{i,\text{adj}} = \begin{cases} 0 & \text{if } i \in I \\ b_i(\alpha'_i) & \text{otherwise.} \end{cases}$

$\alpha'$ ... level of hypothesis $H_i$ in the final graph.
Construction of Adjusted Confidence Bounds

Assumptions:

- Test for $H_i : \theta_i \leq 0 \text{ v.s. } H'_i : \theta_i > 0$
- Let $p_i(\mu)$ denote a p-value for $H_i(\mu) : \theta_i \leq \mu$.
- $p_i(\mu)$ is increasing in $\mu$.
- $b_i(\gamma) = \inf\{\mu | p_i(\mu) > \gamma\}$ (local level $\gamma$ confidence bound)
- $I \subseteq \{1, \ldots, m\}$ ... index set of rejected hypotheses $H_i$.

The adjusted bounds

- If $I = \{1, \ldots, m\}$: $b^{adj}_i = \max\{0, b_i(\alpha_i)\}$.
- Otherwise: $b^{adj}_i = \begin{cases} 0 & \text{if } i \in I \\ b_i(\alpha'_i) & \text{otherwise.} \end{cases}$ where $\alpha'_i$ ... level of hypothesis $H_i$ in the final graph.
Construction of Adjusted Confidence Bounds

Assumptions:

- Test for $H_i : \theta_i \leq 0$ v.s. $H'_i : \theta_i > 0$
- Let $p_i(\mu)$ denote a p-value for $H_i(\mu) : \theta_i \leq \mu$.
- $p_i(\mu)$ is increasing in $\mu$.
- $b_i(\gamma) = \inf\{\mu | p_i(\mu) > \gamma\}$ (local level $\gamma$ confidence bound)
- $I \subseteq \{1, \ldots, m\}$ . . . index set of rejected hypotheses $H_i$.

The adjusted bounds

- If $I = \{1, \ldots, m\}$: $b_i^{adj} = \max\{0, b_i(\alpha_i)\}$.
- Otherwise: $b_i^{adj} = \begin{cases} 0 & \text{if } i \in I \\ b_i(\alpha'_i) & \text{otherwise.} \end{cases}$

$\alpha'$ . . . level of hypothesis $H_i$ in the final graph.
Construction of Adjusted Confidence Bounds

Assumptions:

- Test for $H_i : \theta_i \leq 0$ v.s. $H'_i : \theta_i > 0$
- Let $p_i(\mu)$ denote a p-value for $H_i(\mu) : \theta_i \leq \mu$.
- $p_i(\mu)$ is increasing in $\mu$.
- $b_i(\gamma) = \inf\{\mu | p_i(\mu) > \gamma\}$ (local level $\gamma$ confidence bound)
- $I \subseteq \{1, \ldots, m\}$... index set of rejected hypotheses $H_i$.

The adjusted bounds

- If $I = \{1, \ldots, m\}$: $b_{i}^{adj} = \max\{0, b_i(\alpha_i)\}$.
- Otherwise: $b_{i}^{adj} = \begin{cases} 0 & \text{if } i \in I \\ b_i(\alpha'_i) & \text{otherwise.} \end{cases}$

$\alpha'_i$... level of hypothesis $H_i$ in the final graph.
### Construction of Adjusted Confidence Bounds

**Assumptions:**

- Test for $H_i : \theta_i \leq 0$ v.s. $H'_i : \theta_i > 0$
- Let $p_i(\mu)$ denote a p-value for $H_i(\mu) : \theta_i \leq \mu$.
- $p_i(\mu)$ is increasing in $\mu$.
- $b_i(\gamma) = \inf\{\mu | p_i(\mu) > \gamma\}$ (local level $\gamma$ confidence bound)
- $I \subseteq \{1, \ldots, m\}$ . . . index set of rejected hypotheses $H_i$.

**The adjusted bounds**

- If $I = \{1, \ldots, m\}$: $b_i^{adj} = \max\{0, b_i(\alpha_i)\}$.
- Otherwise:
  
  $b_i^{adj} = \begin{cases} 
  0 & \text{if } i \in I \\
  b_i(\alpha'_i) & \text{otherwise.} 
  \end{cases}$

$\alpha'_i$ . . . level of hypothesis $H_i$ in the final graph.
Construction of Adjusted Confidence Bounds

Assumptions:

• Test for $H_i : \theta_i \leq 0$ v.s. $H_i' : \theta_i > 0$
• Let $p_i(\mu)$ denote a p-value for $H_i(\mu) : \theta_i \leq \mu$.
• $p_i(\mu)$ is increasing in $\mu$.
• $b_i(\gamma) = \inf\{\mu | p_i(\mu) > \gamma\}$ (local level $\gamma$ confidence bound)
• $I \subseteq \{1, \ldots, m\} \ldots$ index set of rejected hypotheses $H_i$.

The adjusted bounds

• If $I = \{1, \ldots, m\}$: $b_i^{adj} = \max\{0, b_i(\alpha_i)\}$.
• Otherwise: $b_i^{adj} = \begin{cases} 0 & \text{if } i \in I \\ b_i(\alpha_i') & \text{otherwise.} \end{cases}$

$\alpha_i' \ldots$ level of hypothesis $H_i$ in the final graph.
Construction of adjusted p-values

Let \( w = (w_1, \ldots, w_m) = (\alpha_1, \ldots, \alpha_m)/\alpha \)

\( J = \{1, \ldots, m\} \) and \( p_{\text{max}} = 0 \)

1. Let \( j = \arg\min_{i \in J} p_i / w_i \)
2. \( p_{j}^{\text{adj}} = \max\{p_j / w_j, p_{\text{max}}\} \)
3. \( p_{\text{max}} = p_{j}^{\text{adj}} \)
4. Update the graph:

\[
J \rightarrow J/\{j\}
\]

\[
w_{\ell} \rightarrow \begin{cases} 
w_{\ell} + w_j g_{j\ell}, & \ell \in J \\
0, & \text{otherwise}
\end{cases}
\]

\[
g_{\ell k} \rightarrow \begin{cases} 
g_{\ell k} + g_{j\ell} g_{j k} / (1 - g_{j\ell} g_{j k}), & \ell, k \in J, \ell \neq k \\
0, & \text{otherwise}
\end{cases}
\]

5. Goto step 1.
Construction of adjusted p-values

Let \( w = (w_1, \ldots, w_m) = (\alpha_1, \ldots, \alpha_m)/\alpha \)

\( J = \{1, \ldots, m\} \) and \( p_{\text{max}} = 0 \)

1. Let \( j = \arg\min_{i \in J} p_i/w_i \)
2. \( p_j^{\text{adj}} = \max\{p_j/w_j, p_{\text{max}}\} \)
3. \( p_{\text{max}} = p_j^{\text{adj}} \)
4. Update the graph:

\[ J \rightarrow J/\{j\} \]

\[ w_\ell \rightarrow \begin{cases} 
  w_\ell + w_j g_{j\ell}, & \ell \in J \\
  0, & \text{otherwise}
\end{cases} \]

\[ g_{\ell k} \rightarrow \begin{cases} 
  \frac{g_{\ell k} + g_{j\ell} g_{j k}}{1 - g_{j\ell} g_{j e}}, & \ell, k \in J, \ell \neq k \\
  0, & \text{otherwise}
\end{cases} \]

5. Goto step 1.
Construction of adjusted p-values

Let \( w = (w_1, \ldots, w_m) = (\alpha_1, \ldots, \alpha_m)/\alpha \)
\( J = \{1, \ldots, m\} \) and \( p_{\text{max}} = 0 \)

1. Let \( j = \arg\min_{i \in J} p_i/w_i \)
2. \( p_{j}^{\text{adj}} = \max\{p_j/w_j, p_{\text{max}}\} \)
3. \( p_{\text{max}} = p_{j}^{\text{adj}} \)
4. Update the graph:

\[
J \rightarrow J/\{j\}
\]

\[
w_\ell \rightarrow \begin{cases} w_\ell + w_j g_{j\ell}, & \ell \in J \\ 0, & \text{otherwise} \end{cases}
\]

\[
g_{\ell k} \rightarrow \begin{cases} 
\frac{g_{\ell k} + g_{j\ell} g_{j k}}{1 - g_{j\ell} g_{j k}}, & \ell, k \in J, \ell \neq k \\ 0, & \text{otherwise} \end{cases}
\]

5. Goto step 1.
Construction of adjusted p-values

Let $\mathbf{w} = (w_1, \ldots, w_m) = (\alpha_1, \ldots, \alpha_m)/\alpha$
$J = \{1, \ldots, m\}$ and $p_{\text{max}} = 0$

1. Let $j = \arg\min_{i \in J} p_i/w_i$
2. $p_j^{\text{adj}} = \max\{p_j/w_j, p_{\text{max}}\}$
3. $p_{\text{max}} = p_j^{\text{adj}}$
4. Update the graph:

$$J \rightarrow J/\{j\}$$

$$w_\ell \rightarrow \begin{cases} w_\ell + w_j g_{j\ell}, & \ell \in J \\ 0, & \text{otherwise} \end{cases}$$

$$g_{\ell k} \rightarrow \begin{cases} \frac{g_{\ell k} + g_{j\ell} g_{j k}}{1 - g_{j\ell} g_{j k}}, & \ell, k \in J, \ell \neq k \\ 0, & \text{otherwise} \end{cases}$$

5. Goto step 1.
Construction of adjusted p-values

Let \( w = (w_1, \ldots, w_m) = (\alpha_1, \ldots, \alpha_m)/\alpha \)
\( J = \{1, \ldots, m\} \) and \( \rho_{\text{max}} = 0 \)

1. Let \( j = \arg\min_{i \in J} p_i/w_i \)
2. \( p_j^{\text{adj}} = \max\{p_j/w_j, \rho_{\text{max}}\} \)
3. \( \rho_{\text{max}} = p_j^{\text{adj}} \)
4. Update the graph:

\[
J \rightarrow J/\{j\}
\]
\[
w_\ell \rightarrow \begin{cases} 
w_\ell + w_j g_\ell, & \ell \in J \\
0, & \text{otherwise} \end{cases}
\]
\[
g_{\ell k} \rightarrow \begin{cases} 
g_{\ell k} + g_j g_{jk} & \ell, k \in J, \ell \neq k \\
1-g_j g_{jk}, & \ell \in J, k \neq j \\
0, & \text{otherwise} \end{cases}
\]
5. Goto step 1.
Example: Improved Fallback Procedure

\[
p_1 = 0.02 \quad p_2 = 0.01 \quad p_3 = 0.06
\]
Example: Improved Fallback Procedure

\[ p_1 = 0.02 \quad \quad p_2 = 0.01 \quad \quad p_3 = 0.06 \]

\[ \frac{p_1}{w_1} = 0.036 \quad \frac{p_2}{w_2} = 0.03 \quad \frac{p_3}{w_3} = 0.36 \]
Example: Improved Fallback Procedure

\[ p_1 = 0.02 \quad p_2 = 0.01 \quad p_3 = 0.06 \]

\[ \frac{p_1}{w_1} = 0.036 \quad p_2^{\text{adj}} = 0.03 \quad \frac{p_3}{w_3} = 0.36 \]
Example: Improved Fallback Procedure

\[
\begin{align*}
      p_1 &= 0.02 & p_2 &= 0.01 & p_3 &= 0.06 \\
\frac{p_1}{w_1} &= 0.024 & p_2^{\text{adj}} &= 0.03 & \frac{p_3}{w_3} &= 0.36
\end{align*}
\]
Example: Improved Fallback Procedure

\[ p_1 = 0.02 \quad \quad p_2 = 0.01 \quad \quad p_3 = 0.06 \]

\[ p_1^{\text{adj}} = 0.03 \quad \quad p_2^{\text{adj}} = 0.03 \quad \quad \frac{p_3}{w_3} = 0.36 \]
Example: Improved Fallback Procedure

$p_1 = 0.02$  $p_2 = 0.01$  $p_3 = 0.06$

$p_1^{adj} = 0.03$  $p_2^{adj} = 0.03$  $\frac{p_3}{w_3} = 0.06$
Example: Improved Fallback Procedure

\[ p_1 = 0.02 \quad p_2 = 0.01 \quad p_3 = 0.06 \]

\[ p_1^{adj} = 0.03 \quad p_2^{adj} = 0.03 \quad p_3^{adj} = 0.06 \]
Case study I
Late phase development of a new drug for the indication of multiple sclerosis

- Two dose levels
- Three hierarchically ordered endpoints: annualized relapse rate, number of lesions in the brain, and disability progression.
- Six elementary hypotheses $H_{ij} : \theta_{ij} \leq 0$
  
  \begin{align*}
  i &= \text{H(igh dose), L(ow dose)} \\
  j &= 1, 2, 3 \ldots \text{endpoints}
  \end{align*}
Strategy 1: Fixed Sequence Test
Strategy 2: Fixed Sequence Test per Dose

\[ \begin{align*}
  H_H1 & \quad \frac{\alpha}{2} \\
  H_H2 & \quad 1 \quad 1 \\
  H_H3 & \quad 1 \\
  H_L1 & \quad \frac{\alpha}{2} \\
  H_L2 & \quad 1 \quad 1 \\
  H_L3 & \quad 1
\end{align*} \]
Strategy 3: More weight to the Primary Endpoints

\[
\begin{align*}
H_{H1} &\rightarrow H_{L1} \quad \frac{\alpha}{2} \\
H_{H2} &\rightarrow H_{L1} \quad \frac{\alpha}{2} \\
H_{H3} &\rightarrow H_{L1} \quad \frac{\alpha}{2} \\
H_{L1} &\rightarrow H_{H1} \quad \frac{1}{2} \\
H_{L1} &\rightarrow H_{H2} \quad \frac{1}{2} \\
H_{L1} &\rightarrow H_{H3} \quad \frac{1}{2} \\
H_{L2} &\rightarrow H_{H2} \quad \frac{1}{2} \\
H_{L2} &\rightarrow H_{H3} \quad \frac{1}{2} \\
H_{L3} &\rightarrow H_{H3} \quad 1 \\
H_{L3} &\rightarrow H_{H1} \quad 1 \\
\end{align*}
\]
Strategy 4: Gatekeeper

\[ H_{H1} \]

\[ H_{H2} \]

\[ H_{H3} \]

\[ H_{L1} \]

\[ H_{L2} \]

\[ H_{L3} \]

\( \alpha \)

\( \frac{1}{2} \)

\( \frac{1}{2} \)

\( 1 - \epsilon \)

\( 1 - \epsilon \)

\( \epsilon \)

\( \epsilon \)

\( 1 - \epsilon \)

\( 1 - \epsilon \)

\( \epsilon \)

\( \epsilon \)
Case Study II
Late phase development of a new cardiovascular drug

- Combination (AB) and mono therapy (B) compared with comparator (A)
- Superiority and non-inferiority tests for primary and multiple secondary endpoints.
- Three elementary hypotheses and two families of hypotheses:
  - $H_1$: superiority of AB vs. A
  - $H_2$: non-inferiority of B vs. A
  - $H_3$: superiority of B vs. A
  - $H_4$: multiple secondary variables for AB vs. A
  - $H_5$: multiple secondary variables for B vs. A
Multiple Test Procedure

\[ H_1 \quad \text{and} \quad H_2 \]
Multiple Test Procedure

$H_1$

$H_2$

$H_3$

$H_4$

$H_5$
Multiple Test Procedure

$\alpha/2$

$H_1$

$\alpha/2$

$H_2$

$H_3$

$H_4$

$H_5$
Multiple Test Procedure

$H_1$ $\alpha/2$

$H_2$ $\alpha/2$

$H_3$ 0

$H_4$ 0

$H_5$ 0
Multiple Test Procedure

$H_1$ \(\alpha/2\) \(\rightarrow\) $H_2$ \(\alpha/2\)

$H_4$ \(1/4\) \(\rightarrow\) $H_3$ \(0\)

$H_5$ \(0\)
Multiple Test Procedure

\[
\begin{align*}
\mathcal{H}_1 & \quad \alpha/2 \\
\mathcal{H}_2 & \quad \alpha/2 \\
\mathcal{H}_3 & \quad 0 \\
\mathcal{H}_4 & \quad (0, 1/4) \\
\mathcal{H}_5 & \quad (0, 3/4)
\end{align*}
\]
Multiple Test Procedure

\[ H_1 \xrightarrow{\alpha/2} H_2 \]
\[ H_2 \xrightarrow{\alpha/2} H_3 \]
\[ H_3 \xrightarrow{1/4} H_4 \]
\[ H_3 \xrightarrow{3/4} H_5 \]

\[ H_1 \xrightarrow{3/4} H_2 \]
\[ H_2 \xrightarrow{1/4} H_3 \]

\[ H_4 \]
\[ H_5 \]
Multiple Test Procedure
Multiple Test Procedure

\[ \alpha = \left( \frac{\alpha}{2}, \frac{\alpha}{2}, 0, 0, 0 \right) \]

\[ G = \begin{pmatrix}
0 & 3/4 & 0 & 1/4 & 0 \\
0 & 0 & 3/4 & 0 & 1/4 \\
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
\end{pmatrix} \]
Example
Example
Example
Example
Example
Example

\[ \frac{1}{2} \alpha \quad \mathcal{H}_4 \quad \frac{1}{2} \alpha \]

\[ 1 \quad \mathcal{H}_5 \quad 1 \]
Summary and Extensions

• Intuitive graphical procedure to construct multiple tests
• Easy to communicate the testing strategy
• Easy to implement in software
• Adjusted p-values available
• Multiplicity adjusted confidence intervals can be constructed based on Strassburger and Bretz (2008), Guilbaud (2008)
• Adjusted p-values
• Interpretation as Finite Markov Chain
• Similar approach published by Burman (2009)
Selected References

P. Bauer, W. Brannath, and M. Posch.
Multiple testing for identifying effective and safe treatments.

F. Bretz, W. Maurer, W. Brannath, and M. Posch.
A graphical approach to sequentially rejective multiple test procedures.

A. Dmitrienko, W.W. Offen, and P.H. Westfall.
Gatekeeping strategies for clinical trials that do not require all primary effects to be significant.

O. Guilbaud.
Simultaneous confidence regions corresponding to holm’s stepdown procedure and other closed-testing procedures.

G. Hommel, F. Bretz, and W. Maurer.
Powerful short-cuts for multiple testing procedures with special reference to gatekeeping strategies.

K. Strassburger and F. Bretz.
Compatible simultaneous lower confidence bounds for the holm procedure and other bonferroni based closed tests.