Part 2

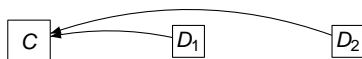## Graphical Approaches to Multiple Testing

Structured families of hypotheses

Example for structured hypotheses

Example
A parallel group study with

- Two treatments and a control

- One primary and one secondary endpoint
    - For example, FEV1 and time to exacerbation in a COPD trial (see Part 1)

Structure:

- Four hypotheses of interest
    - Two are primary (FEV1 for low and high dose)
    - Two secondary (time to exacerbation for low and high dose)

Precise role of study objectives

Key question in case of several study objectives/hypotheses:

What is their precise role?

Primary?		Required for study success?
Secondary?	⇔	For additional label claims?
Tertiary?		Just exploratory?

- Type I error rate control may be only required for some (maybe just one primary) hypothesis
- Classification is specific to the study, needs discussion within clinical teams and with regulatory agencies

Precise role of primary hypothesis

The primary hypothesis may be

- Clinically more important than the secondary hypothesis
  - Example: Glucose level in blood (e.g. HbA1c) and weight loss in diabetes
- Key to approval, although not clinically more important
  - Example: FEV1 and time to exacerbation in COPD
- Prerequisite for possible significance in the secondary hypothesis
  - Example: non-inferiority needs to be established before testing superiority

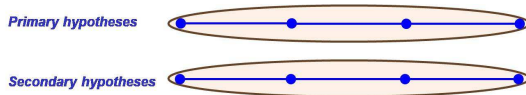# Structured hypotheses

Traditional multiple testing methods assume

- No order (all hypotheses equally important)
- Strict hierarchy (e.g. H_2 only tested if H_1 is rejected before)
- Strict hierarchy in blocks:

Primary hypotheses

Secondary hypotheses
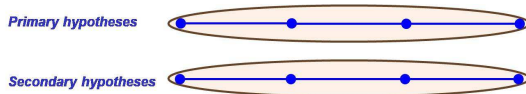
Structured hypotheses

Traditional multiple testing methods assume

- No order (all hypotheses equally important)
- Strict hierarchy (e.g. H_2 only tested if H_1 is rejected before)
- Strict hierarchy in blocks:

Concrete applications often impose more structure through
successiveness property between primary and corresponding
secondary hypotheses:
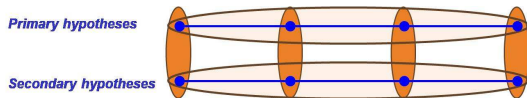
Structured hypotheses – example revisited

- The two doses are equally relevant
- FEV1 increase and reduction in time to exacerbation within a dose group are successive
    - That is, for the same dose, time to exacerbation is only of interest if FEV1 increase has been shown before
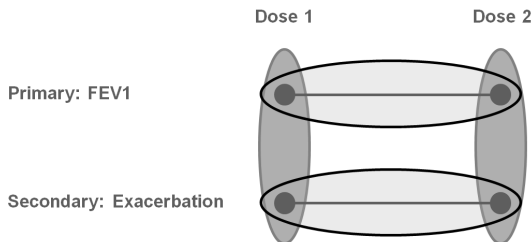
How to construct decision strategies
that reflect such complex requirements?

Standard multiple comparison procedures ...

- include Bonferroni, Holm, Hochberg, Dunnett, etc.
- control the FWER at level α
- are not suitable, because they treat all hypotheses equally and do not address the underlying structure of the test problem

Need for suitable multiple test procedures

Standard multiple comparison procedures ...

- include Bonferroni, Holm, Hochberg, Dunnett, etc.
- control the FWER at level α
- are not suitable, because they treat all hypotheses equally and do not address the underlying structure of the test problem

An intuitive procedure that does not control the FWER

- Test H_1, H_2 with Holm at level α; if at least one is rejected, test the "descendant" secondary hypothesis at level $\alpha/2$.
- This procedure (or many variants thereof) does not control the FWER at level α; actual error rate can be up to $3\alpha/2$.

- Often applied to clinical trials with structured families of hypotheses and several levels of multiplicity
  - Multiple endpoints (e.g. HbA1c and body weight)
  - Multiple treatments of the same drug (e.g. dose or regimen)
  - Multiple populations (e.g. full and sub-population)
  - Combined non-inferiority and superiority testing
- Reflect the difference in importance as well as the relationship between the various study objectives
- Can be represented as weighted closed test procedures

- Often applied to clinical trials with structured families of hypotheses and several levels of multiplicity
  - Multiple endpoints (e.g. HbA1c and body weight)
  - Multiple treatments of the same drug (e.g. dose or regimen)
  - Multiple populations (e.g. full and sub-population)
  - Combined non-inferiority and superiority testing
- Reflect the difference in importance as well as the relationship between the various study objectives
- Can be represented as weighted closed test procedures

- Often applied to clinical trials with structured families of hypotheses and several levels of multiplicity
  - Multiple endpoints (e.g. HbA1c and body weight)
  - Multiple treatments of the same drug (e.g. dose or regimen)
  - Multiple populations (e.g. full and sub-population)
  - Combined non-inferiority and superiority testing

- Reflect the difference in importance as well as the relationship between the various study objectives

- Can be represented as weighted closed test procedures

- Often applied to clinical trials with structured families of hypotheses and several levels of multiplicity
  - Multiple endpoints (e.g. HbA1c and body weight)
  - Multiple treatments of the same drug (e.g. dose or regimen)
  - Multiple populations (e.g. full and sub-population)
  - Combined non-inferiority and superiority testing

- Reflect the difference in importance as well as the relationship between the various study objectives

- Can be represented as weighted closed test procedures

Closed test procedure

Hypotheses:

- $H_1, \ldots, H_m$: m elementary null hypotheses
- $2^m - 1$ intersection hypotheses

  $H_J = H_{j_1} \cap H_{j_2} \cap \ldots \cap H_{j_k}, J = \{j_1, \ldots, j_k\} \subseteq \{1, \ldots, m\}$

Closed test procedure (CTP):

- Test each H_J with a suitable α-level test
- Reject an elementary hypothesis H_j, if all intersection hypotheses containing the index j can be rejected
- This controls the FWER strongly at level α

Weighted Bonferroni test

- Test $H_1, \ldots, H_m$ at level α with weights $w_1, \ldots, w_m \geq 0$ such that $w_1 + \ldots + w_m \leq 1$, i.e. $w_1 \alpha + \ldots + w_m \alpha \leq \alpha$

  - Assume m unadjusted p-values $p_1, \ldots, p_m$

  - Weighted Bonferroni test (for the global null hypothesis):

    Reject $H = H_1 \cap \ldots \cap H_m$ if $p_j \leq w_j \alpha$ for at least one j

  - Alternatively, define weighted p-values $\tilde{p}_j = p_j / w_j$ and reject H if $\min_j \tilde{p}_j \leq \alpha$

  - If $w_j = 1/m$, this results in the ordinary, unweighted Bonferroni test

  - CTP using weighted Bonferroni tests for each intersection hypothesis H_J controls the FWER strongly at level α

  - See below on how to select weights $w_j(J)$ for each J

Weighted Bonferroni test

- Test $H_1, \ldots, H_m$ at level α with weights $w_1, \ldots, w_m \geq 0$ such that $w_1 + \ldots + w_m \leq 1$, i.e. $w_1 \alpha + \ldots + w_m \alpha \leq \alpha$
- Assume m unadjusted p-values $p_1, \ldots, p_m$
- Weighted Bonferroni test (for the global null hypothesis): Reject $H = H_1 \cap \ldots \cap H_m$ if $p_j \leq w_j \alpha$ for at least one j
- Alternatively, define weighted p-values $\tilde{p}_j = p_j / w_j$ and reject H if $\min_j \tilde{p}_j \leq \alpha$
- If $w_j = 1/m$, this results in the ordinary, unweighted Bonferroni test
- CTP using weighted Bonferroni tests for each intersection hypothesis H_J controls the FWER strongly at level α
- See below on how to select weights $w_j(J)$ for each J

Weighted Bonferroni test

- Test $H_1, \ldots, H_m$ at level α with weights $w_1, \ldots, w_m \geq 0$ such that $w_1 + \ldots + w_m \leq 1$, i.e. $w_1\alpha + \ldots + w_m\alpha \leq \alpha$
- Assume m unadjusted p-values $p_1, \ldots, p_m$
- Weighted Bonferroni test (for the global null hypothesis):

    Reject $H = H_1 \cap \ldots \cap H_m$ if $p_j \leq w_j\alpha$ for at least one j

- Alternatively, define weighted p-values $\tilde{p}_j = p_j/w_j$ and reject H if $\min_j \tilde{p}_j \leq \alpha$
- If $w_j = 1/m$, this results in the ordinary, unweighted Bonferroni test
- CTP using weighted Bonferroni tests for each intersection hypothesis H_J controls the FWER strongly at level α
- See below on how to select weights $w_j(J)$ for each J

Weighted Bonferroni test

- Test $H_1, \ldots, H_m$ at level α with weights $w_1, \ldots, w_m \geq 0$ such that $w_1 + \ldots + w_m \leq 1$, i.e. $w_1\alpha + \ldots + w_m\alpha \leq \alpha$
- Assume m unadjusted p-values $p_1, \ldots, p_m$
- Weighted Bonferroni test (for the global null hypothesis):
    Reject $H = H_1 \cap \ldots \cap H_m$ if $p_j \leq w_j\alpha$ for at least one j
- Alternatively, define weighted p-values $\tilde{p}_j = p_j/w_j$ and reject H if $\min_j \tilde{p}_j \leq \alpha$
- If $w_i = 1/m$, this results in the ordinary, unweighted Bonferroni test
- CTP using weighted Bonferroni tests for each intersection hypothesis H_J controls the FWER strongly at level α
- See below on how to select weights $w_j(J)$ for each J

Weighted Bonferroni test

- Test $H_1, \ldots, H_m$ at level α with weights $w_1, \ldots, w_m \geq 0$ such that $w_1 + \ldots + w_m \leq 1$, i.e. $w_1\alpha + \ldots + w_m\alpha \leq \alpha$
- Assume m unadjusted p-values $p_1, \ldots, p_m$
- Weighted Bonferroni test (for the global null hypothesis):
  Reject $H = H_1 \cap \ldots \cap H_m$ if $p_j \leq w_j\alpha$ for at least one j
- Alternatively, define weighted p-values $\tilde{p}_j = p_j/w_j$ and reject H if $\min_j \tilde{p}_j \leq \alpha$
- If $w_j = 1/m$, this results in the ordinary, unweighted Bonferroni test
- CTP using weighted Bonferroni tests for each intersection hypothesis H_J controls the FWER strongly at level α
- See below on how to select weights $w_j(J)$ for each J

Weighted Bonferroni test

- Test $H_1, \ldots, H_m$ at level α with weights $w_1, \ldots, w_m \geq 0$ such that $w_1 + \ldots + w_m \leq 1$, i.e. $w_1\alpha + \ldots + w_m\alpha \leq \alpha$
- Assume m unadjusted p-values $p_1, \ldots, p_m$
- Weighted Bonferroni test (for the global null hypothesis):

  Reject $H = H_1 \cap \ldots \cap H_m$ if $p_j \leq w_j\alpha$ for at least one j
- Alternatively, define weighted p-values $\tilde{p}_j = p_j/w_j$ and reject H if $\min_j \tilde{p}_j \leq \alpha$
- If $w_j = 1/m$, this results in the ordinary, unweighted Bonferroni test
- CTP using weighted Bonferroni tests for each intersection hypothesis H_J controls the FWER strongly at level α
- See below on how to select weights $w_j(J)$ for each J

Weighted Bonferroni test

- Test $H_1, \ldots, H_m$ at level α with weights $w_1, \ldots, w_m \geq 0$ such that $w_1 + \ldots + w_m \leq 1$, i.e. $w_1\alpha + \ldots + w_m\alpha \leq \alpha$
- Assume m unadjusted p-values $p_1, \ldots, p_m$
- Weighted Bonferroni test (for the global null hypothesis):
  Reject $H = H_1 \cap \ldots \cap H_m$ if $p_j \leq w_j\alpha$ for at least one j
- Alternatively, define weighted p-values $\tilde{p}_j = p_j/w_j$ and reject H if $\min_j \tilde{p}_j \leq \alpha$
- If $w_j = 1/m$, this results in the ordinary, unweighted Bonferroni test
- CTP using weighted Bonferroni tests for each intersection hypothesis H_J controls the FWER strongly at level α
- See below on how to select weights $w_j(J)$ for each J

Consonance

- CTP is consonant, if the following condition is satisfied:
  If $H = H_1 \cap \ldots \cap H_m$ is rejected, then reject at least one H_j

- Desirable property of CTP, as it ensures rejection of an individual null after rejecting the global null

- Consonance enables construction of sequentially rejective procedures, reducing the number of tests from $2^m - 1$ to m

- Not all CTP are consonant
  - Assume $m = 2$ and test H_1, H_2 using a CTP
  - If $H_{12} = H_1 \cap H_2$ is rejected with a weighted Bonferroni test (i.e. either $p_1 \leq w_1 \alpha$ or $p_2 \leq w_2 \alpha$), then either H_1 or H_2 can be rejected at α, respectively (consonance).
  - If H_{12} is rejected with an F test, it is possible that neither H_1 nor H_2 is rejected at α (inconsonance).

Consonance

- CTP is consonant, if the following condition is satisfied:

  If $H = H_1 \cap \ldots \cap H_m$ is rejected, then reject at least one H_j

- Desirable property of CTP, as it ensures rejection of an individual null after rejecting the global null

- Consonance enables construction of sequentially rejective procedures, reducing the number of tests from $2^m - 1$ to m

- Not all CTP are consonant
  - Assume $m = 2$ and test H_1, H_2 using a CTP
  - If $H_{12} = H_1 \cap H_2$ is rejected with a weighted Bonferroni test (i.e. either $p_1 \leq w_1\alpha$ or $p_2 \leq w_2\alpha$), then either H_1 or H_2 can be rejected at α, respectively (consonance).
  - If H_{12} is rejected with an F test, it is possible that neither H_1 nor H_2 is rejected at α (inconsonance).

Consonance

- CTP is consonant, if the following condition is satisfied:
  If $H = H_1 \cap \ldots \cap H_m$ is rejected, then reject at least one H_j
- Desirable property of CTP, as it ensures rejection of an individual null after rejecting the global null
- Consonance enables construction of sequentially rejective procedures, reducing the number of tests from $2^m - 1$ to m
- Not all CTP are consonant
  - Assume $m = 2$ and test H_1, H_2 using a CTP
  - If $H_{12} = H_1 \cap H_2$ is rejected with a weighted Bonferroni test (i.e. either $p_1 \leq w_1 \alpha$ or $p_2 \leq w_2 \alpha$), then either H_1 or H_2 can be rejected at α, respectively (consonance).
  - If H_{12} is rejected with an F test, it is possible that neither H_1 nor H_2 is rejected at α (inconsonance).

Consonance

- CTP is consonant, if the following condition is satisfied:

  If $H = H_1 \cap \ldots \cap H_m$ is rejected, then reject at least one H_j

- Desirable property of CTP, as it ensures rejection of an individual null after rejecting the global null

- Consonance enables construction of sequentially rejective procedures, reducing the number of tests from $2^m - 1$ to m

- Not all CTP are consonant
  - Assume $m = 2$ and test H_1, H_2 using a CTP
    - If $H_{12} = H_1 \cap H_2$ is rejected with a weighted Bonferroni test (i.e. either $p_1 \leq w_1\alpha$ or $p_2 \leq w_2\alpha$), then either H_1 or H_2 can be rejected at α, respectively (consonance).
    - If H_{12} is rejected with an F test, it is possible that neither H_1 nor H_2 is rejected at α (inconsonance).

Consonance

- CTP is consonant, if the following condition is satisfied:

  If $H = H_1 \cap \ldots \cap H_m$ is rejected, then reject at least one H_j

- Desirable property of CTP, as it ensures rejection of an individual null after rejecting the global null

- Consonance enables construction of sequentially rejective procedures, reducing the number of tests from $2^m - 1$ to m

- Not all CTP are consonant
  - Assume $m = 2$ and test H_1, H_2 using a CTP
  - If $H_{12} = H_1 \cap H_2$ is rejected with a weighted Bonferroni test (i.e. either $p_1 \leq w_1 \alpha$ or $p_2 \leq w_2 \alpha$), then either H_1 or H_2 can be rejected at α, respectively (consonance).
  - If H_{12} is rejected with an F test, it is possible that neither H_1 nor H_2 is rejected at α (inconsonance).

Consonance

- CTP is consonant, if the following condition is satisfied:

  If $H = H_1 \cap \ldots \cap H_m$ is rejected, then reject at least one H_j

- Desirable property of CTP, as it ensures rejection of an individual null after rejecting the global null

- Consonance enables construction of sequentially rejective procedures, reducing the number of tests from $2^m - 1$ to m

- Not all CTP are consonant
  - Assume $m = 2$ and test H_1, H_2 using a CTP
  - If $H_{12} = H_1 \cap H_2$ is rejected with a weighted Bonferroni test (i.e. either $p_1 \leq w_1\alpha$ or $p_2 \leq w_2\alpha$), then either H_1 or H_2 can be rejected at α, respectively (consonance).
  - If H_{12} is rejected with an F test, it is possible that neither H_1 nor H_2 is rejected at α (inconsonance).

- For each $J \subseteq \{1, \ldots, m\}$ choose weights $w_j(J) \geq 0$ such that $\sum_j w_j(J) \leq 1$

- Reject H_j, if $p_j \leq w_j(J)\alpha$ for at least one $j \in J$

- If for every pair $I, J \subseteq \{1, \ldots m\}$ with $J \subseteq I$ the monotonicity

  $$w_j(I) \leq w_j(J) \text{ for all } j \in J$$

  holds, then consonance is satisfied

- Resulting CTP can be performed as sequentially rejective procedure based on weighted p-values $\tilde{p}_j(J) = p_j / w_j(J)$, starting with $J = \{1, \ldots, m\}$:

- Holds for weighted Holm procedure, fixed sequence test, Bonferroni-based fallback and gatekeeping procedures, etc

- For each $J \subseteq \{1, \ldots, m\}$ choose weights $w_j(J) \geq 0$ such that $\sum_j w_j(J) \leq 1$
- Reject H_J, if $p_j \leq w_j(J)\alpha$ for at least one $j \in J$
- If for every pair $I, J \subseteq \{1, \ldots, m\}$ with $J \subseteq I$ the monotonicity

$$w_j(I) \leq w_j(J) \text{ for all } j \in J$$

holds, then consonance is satisfied
- Resulting CTP can be performed as sequentially rejective procedure based on weighted p-values $\tilde{p}_j(J) = p_j / w_j(J)$, starting with $J = \{1, \ldots, m\}$:

- Holds for weighted Holm procedure, fixed sequence test, Bonferroni-based fallback and gatekeeping procedures, etc

Consonance for Closed Weighted Bonferroni Tests

- For each $J \subseteq \{1, \ldots, m\}$ choose weights $w_j(J) \geq 0$ such that $\sum_j w_j(J) \leq 1$
- Reject H_J, if $p_j \leq w_j(J)\alpha$ for at least one $j \in J$
- If for every pair $I, J \subseteq \{1, \ldots m\}$ with $J \subseteq I$ the monotonicity

$$w_j(I) \leq w_j(J) \text{ for all } j \in J$$

holds, then consonance is satisfied

- Resulting CTP can be performed as sequentially rejective procedure based on weighted p-values $\tilde{p}_j(J) = p_j / w_j(J)$, starting with $J = \{1, \ldots, m\}$:

- Holds for weighted Holm procedure, fixed sequence test, Bonferroni-based fallback and gatekeeping procedures, etc

- For each $J \subseteq \{1, \ldots, m\}$ choose weights $w_j(J) \geq 0$ such that $\sum_j w_j(J) \leq 1$
- Reject H_J, if $p_j \leq w_j(J)\alpha$ for at least one $j \in J$
- If for every pair $I, J \subseteq \{1, \ldots m\}$ with $J \subseteq I$ the monotonicity

$$w_j(I) \leq w_j(J) \text{ for all } j \in J$$

  holds, then consonance is satisfied
- Resulting CTP can be performed as sequentially rejective procedure based on weighted p-values $\tilde{p}_j(J) = p_j/w_j(J)$, starting with $J = \{1, \ldots, m\}$:

Holds for weighted Holm procedure, fixed sequence test, Bonferroni-based fallback and gatekeeping procedures, etc

Consonance for Closed Weighted Bonferroni Tests

- For each $J \subseteq \{1, \ldots, m\}$ choose weights $w_j(J) \geq 0$ such that $\sum_j w_j(J) \leq 1$
- Reject H_J, if $p_j \leq w_j(J)\alpha$ for at least one $j \in J$
- If for every pair $I, J \subseteq \{1, \ldots m\}$ with $J \subseteq I$ the monotonicity

$$w_j(I) \leq w_j(J) \text{ for all } j \in J$$

holds, then consonance is satisfied
- Resulting CTP can be performed as sequentially rejective procedure based on weighted p-values $\tilde{p}_j(J) = p_j / w_j(J)$, starting with $J = \{1, \ldots, m\}$:

  1 If $\tilde{p}_i(J) = \min\{\tilde{p}_j(J), j \in J\} \leq \alpha$, reject H_i; otherwise stop.

  2 $J \to J/\{i\}$

  3 Go to step 1.

- Holds for weighted Holm procedure, fixed sequence test, Bonferroni-based fallback and gatekeeping procedures, etc

Consonance for Closed Weighted Bonferroni Tests

- For each $J \subseteq \{1, \ldots, m\}$ choose weights $w_j(J) \geq 0$ such that $\sum_j w_j(J) \leq 1$
- Reject H_J, if $p_j \leq w_j(J)\alpha$ for at least one $j \in J$
- If for every pair $I, J \subseteq \{1, \ldots m\}$ with $J \subseteq I$ the monotonicity

$$w_j(I) \leq w_j(J) \text{ for all } j \in J$$

  holds, then consonance is satisfied
- Resulting CTP can be performed as sequentially rejective procedure based on weighted p-values $\tilde{p}_j(J) = p_j / w_j(J)$, starting with $J = \{1, \ldots, m\}$:
    1. If $\tilde{p}_i(J) = \min\{\tilde{p}_j(J), j \in J\} \leq \alpha$, reject H_i; otherwise stop.
    2. $J \rightarrow J/\{i\}$
    3. Go to step 1.

- Holds for weighted Holm procedure, fixed sequence test, Bonferroni-based fallback and gatekeeping procedures, etc

Consonance for Closed Weighted Bonferroni Tests

- For each $J \subseteq \{1, \ldots, m\}$ choose weights $w_j(J) \geq 0$ such that $\sum_j w_j(J) \leq 1$
- Reject H_J, if $p_j \leq w_j(J)\alpha$ for at least one $j \in J$
- If for every pair $I, J \subseteq \{1, \ldots m\}$ with $J \subseteq I$ the monotonicity

$$w_j(I) \leq w_j(J) \text{ for all } j \in J$$

  holds, then consonance is satisfied
- Resulting CTP can be performed as sequentially rejective procedure based on weighted p-values $\tilde{p}_j(J) = p_j / w_j(J)$, starting with $J = \{1, \ldots, m\}$:
  1. If $\tilde{p}_i(J) = \min\{\tilde{p}_j(J), j \in J\} \leq \alpha$, reject H_i; otherwise stop.
  2. $J \to J/\{i\}$
  3. Go to step 1.
- Holds for weighted Holm procedure, fixed sequence test, Bonferroni-based fallback and gatekeeping procedures, etc

Consonance for Closed Weighted Bonferroni Tests

- For each $J \subseteq \{1, \ldots, m\}$ choose weights $w_j(J) \geq 0$ such that $\sum_j w_j(J) \leq 1$
- Reject H_J, if $p_j \leq w_j(J)\alpha$ for at least one $j \in J$
- If for every pair $I, J \subseteq \{1, \ldots m\}$ with $J \subseteq I$ the monotonicity

$$w_j(I) \leq w_j(J) \text{ for all } j \in J$$

  holds, then consonance is satisfied
- Resulting CTP can be performed as sequentially rejective procedure based on weighted p-values $\tilde{p}_j(J) = p_j / w_j(J)$, starting with $J = \{1, \ldots, m\}$:
  1. If $\tilde{p}_i(J) = \min\{\tilde{p}_j(J), j \in J\} \leq \alpha$, reject H_i; otherwise stop.
  2. $J \rightarrow J/\{i\}$
  3. Go to step 1.
- Holds for weighted Holm procedure, fixed sequence test, Bonferroni-based fallback and gatekeeping procedures, etc

Consonance for Closed Weighted Bonferroni Tests

- For each $J \subseteq \{1, \ldots, m\}$ choose weights $w_j(J) \geq 0$ such that $\sum_j w_j(J) \leq 1$
- Reject H_J, if $p_j \leq w_j(J)\alpha$ for at least one $j \in J$
- If for every pair $I, J \subseteq \{1, \ldots m\}$ with $J \subseteq I$ the monotonicity

$$w_j(I) \leq w_j(J) \text{ for all } j \in J$$

  holds, then consonance is satisfied
- Resulting CTP can be performed as sequentially rejective procedure based on weighted p-values $\tilde{p}_j(J) = p_j / w_j(J)$, starting with $J = \{1, \ldots, m\}$:
  1. If $\tilde{p}_i(J) = \min\{\tilde{p}_j(J), j \in J\} \leq \alpha$, reject H_i; otherwise stop.
  2. $J \to J/\{i\}$
  3. Go to step 1.
- Holds for weighted Holm procedure, fixed sequence test, Bonferroni-based fallback and gatekeeping procedures, etc

Graphical Approach

Heuristics

Notation

- Null hypotheses $H_1, \ldots, H_m$
- Initial allocation of the significance level $\alpha = \alpha_1 + \ldots + \alpha_m$
- Unadjusted p-values $p_1, \ldots, p_m$

Heuristics

Notation

- Null hypotheses $H_1, \ldots, H_m$
- Initial allocation of the significance level $\alpha = \alpha_1 + \ldots + \alpha_m$
- Unadjusted p-values $p_1, \ldots, p_m$

"α propagation"

If a hypothesis H_i can be rejected at level α_i (i.e. $p_i \leq \alpha_i$), reallocate its level α_i to the remaining, not yet rejected hypotheses (according to a prefixed rule) and continue testing with the updated α levels.

1. Hypotheses $H_1, \ldots, H_m$ represented as nodes

(H_1) (H_2)

Conventions

1. Hypotheses $H_1, \ldots, H_m$ represented as nodes

(H_1) (H_2)

2. Split of significance level α as weights $\alpha_1, \ldots, \alpha_m$

$\alpha_1 = \frac{\alpha}{2}$ $\alpha_2 = \frac{\alpha}{2}$

(H_1) (H_2)

Conventions

1. Hypotheses $H_1, \ldots, H_m$ represented as nodes

(H_1) (H_2)

2. Split of significance level α as weights $\alpha_1, \ldots, \alpha_m$

$\frac{\alpha}{2}$ $\frac{\alpha}{2}$
(H_1) (H_2)

3. "α propagation" through weighted, directed edges

Conventions

1. Hypotheses $H_1, \ldots, H_m$ represented as nodes

(H_1) (H_2)

2. Split of significance level α as weights $\alpha_1, \ldots, \alpha_m$

Bonferroni $(H_1)^{\frac{\alpha}{2}}$ $(H_2)^{\frac{\alpha}{2}}$

3. "α propagation" through weighted, directed edges
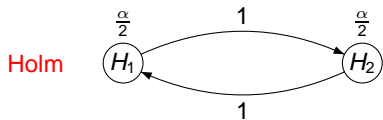
Holm

$$\frac{\alpha}{2}$$

(H_1)

$$\frac{\alpha}{2}$$

(H_2)

Remarks

- Single-step procedures (e.g. Bonferroni) have no α propagation (i.e. no edges between nodes)
- Stepwise procedures (e.g. Holm) include α propagation and are thus more powerful

$\frac{\alpha}{2} = 0.0125$ 1 $\frac{\alpha}{2} = 0.0125$

H_1 H_2

$p_1 = 0.04$ 1 $p_2 = 0.01$

$\alpha = 0.025$

$\left(H_1 \right)$

$p_1 = 0.04$

$\alpha = 0.025$

H_1
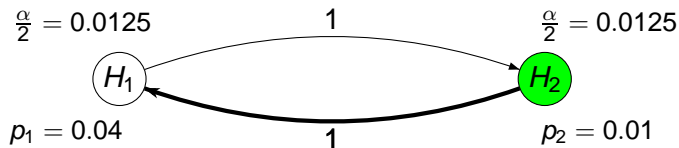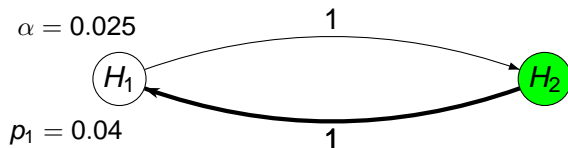
$p_1 = 0.04$

Weighted Bonferroni-Holm test
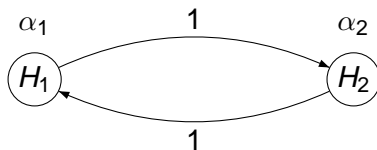
Use α_1, α_2 with $\alpha_1 + \alpha_2 = \alpha$ instead of $\alpha_1 = \alpha_2 = \alpha/2$

Weighted Bonferroni-Holm test

Use α_1, α_2 with $\alpha_1 + \alpha_2 = \alpha$ instead of $\alpha_1 = \alpha_2 = \alpha/2$

Fixed sequence test

$m = 2:$

$m = 3:$
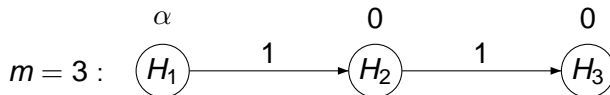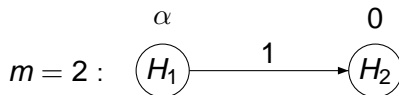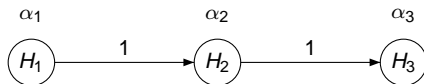
Original fallback

Fallback procedures ($m = 3$)
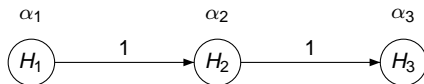
Original fallback

Improved fallback I
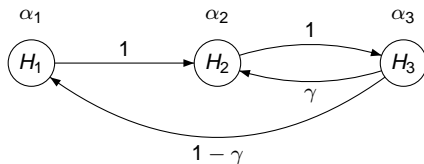
$$\gamma = \frac{\alpha_2}{\alpha_1 + \alpha_2}$$

Fallback procedures ($m = 3$)

Original fallback
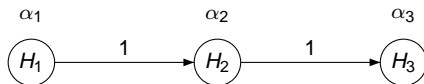
Improved fallback I
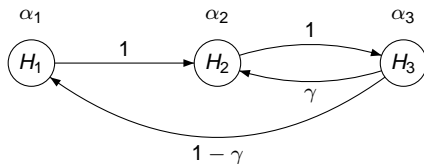
$$\gamma = \frac{\alpha_2}{\alpha_1 + \alpha_2}$$

Improved fallback II

$$\epsilon \to 0$$
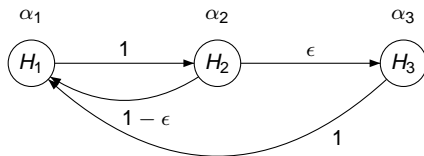
Table I. Weights assigned to the intersection hypothesis tests.

Intersection hypothesis	Weights			
	H_1	H_2	H_3	H_4
$H_1 \cap H_2 \cap H_3 \cap H_4$	0.5	0.5	0.0	0.0
$H_1 \cap H_2 \cap H_3$	0.5	0.5	0.0	0.0
$H_1 \cap H_2 \cap H_4$	0.5	0.5	0.0	0.0
$H_1 \cap H_2$	0.5	0.5	0.0	0.0
$H_1 \cap H_3 \cap H_4$	0.5	0.0	0.25	0.25
$H_1 \cap H_3$	0.5	0.0	0.5	0.0
$H_1 \cap H_4$	0.5	0.0	0.0	0.5
H_1	0.5	0.0	0.0	0.0
$H_2 \cap H_3 \cap H_4$	0.0	0.5	0.25	0.25
$H_2 \cap H_3$	0.0	0.5	0.5	0.0
$H_2 \cap H_4$	0.0	0.5	0.0	0.5
H_2	0.0	0.5	0.0	0.0
$H_3 \cap H_4$	0.0	0.0	0.5	0.5
H_3	0.0	0.0	1.0	0.0
H_4	0.0	0.0	0.0	1.0

(Dmitrienko, Offen & Westfall, 2003)
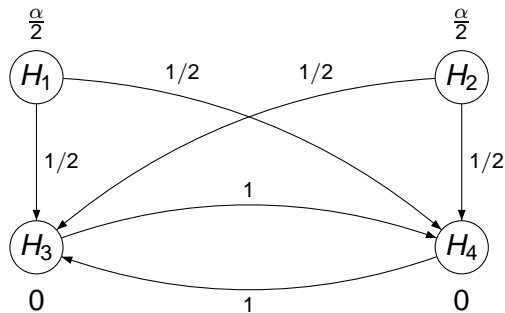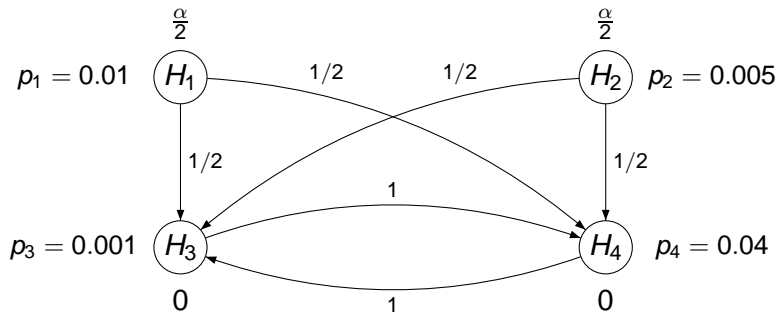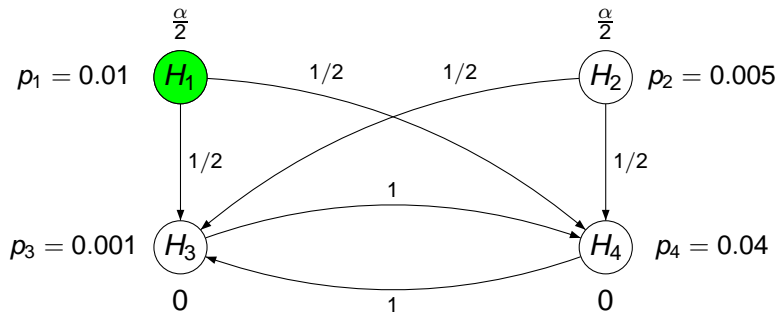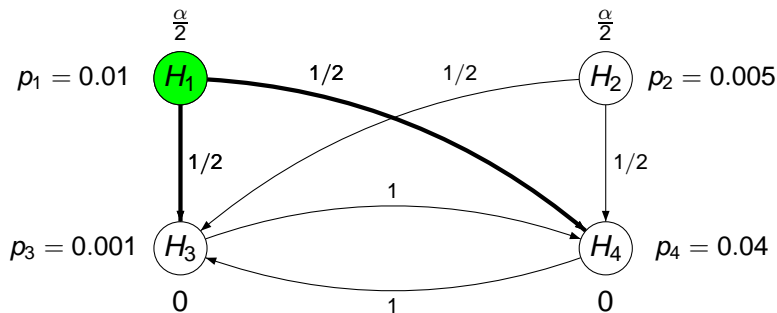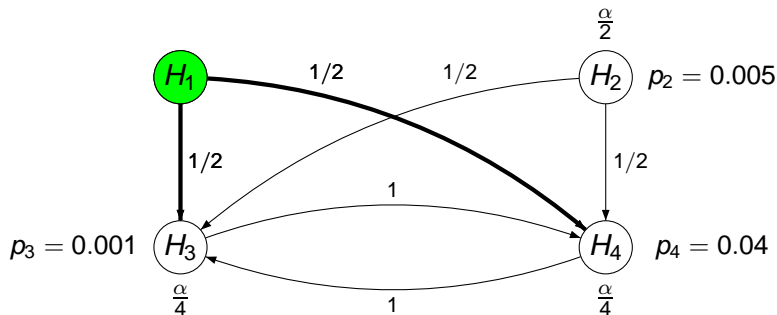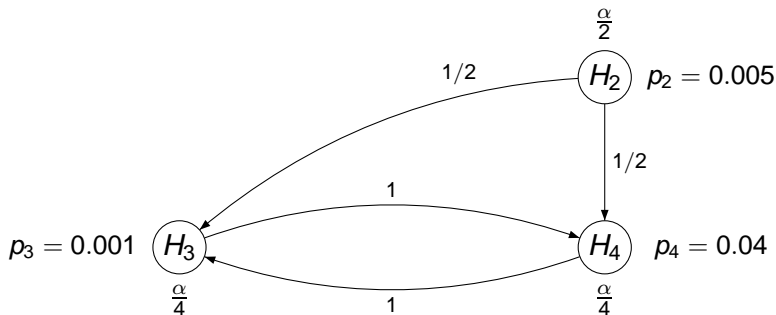
Procedure not successive:

H_4 could be rejected without having H_2 rejected
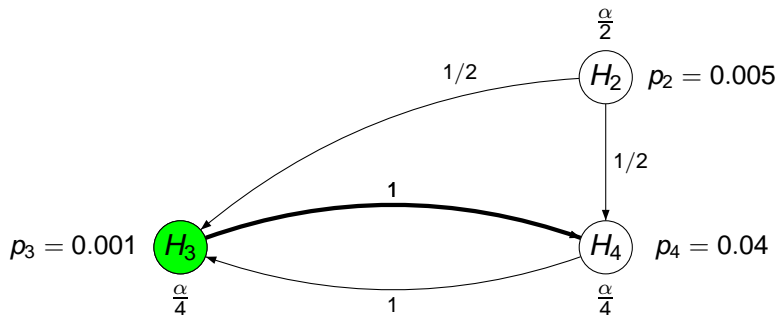
$\frac{\alpha}{2}$

H_2 $p_2 = 0.005$

1

H_4 $p_4 = 0.04$

$\frac{\alpha}{2}$

H_4 $p_4 = 0.04$

α
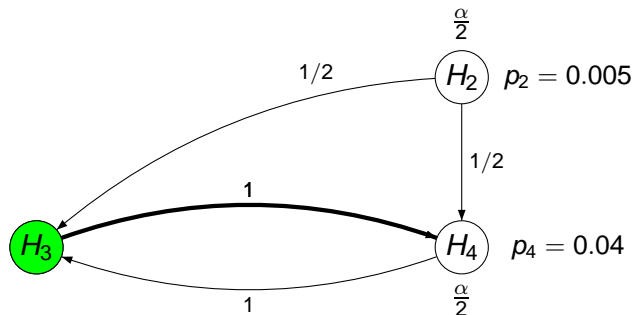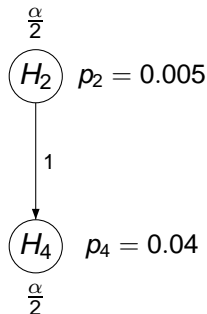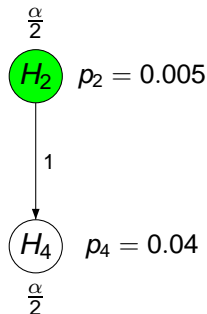
Formal definition of the graphical approach

General Definition

- $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$, $\sum_{i=1}^{m} \alpha_i = \alpha$, initial levels
- $\mathbf{G} = (g_{ij})$: $m \times m$ 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 propagated to H_j
- $\mathbf{G}$ and α determine the graph and the multiple test

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:

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

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

$$g_{\ell m} \rightarrow \begin{cases} \frac{g_{\ell m} + g_{\ell j} g_{jm}}{1 - g_{\ell j} g_{j\ell}}, & \ell, m \in J, \ell \neq m, g_{\ell j} g_{j\ell} < 1 \\ 0, & \text{otherwise} \end{cases}$$
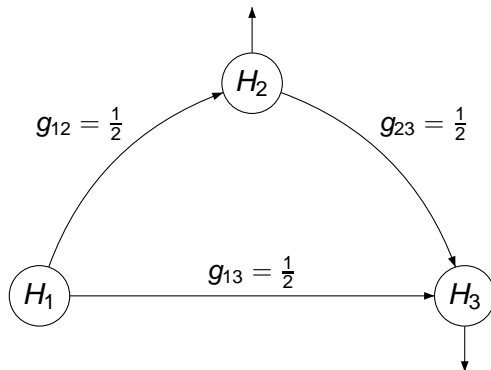
③ Go to step 1.

Update algorithm

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:

$$J \to J/\{j\}$$

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

$$g_{\ell m} \to \begin{cases} \frac{g_{\ell m} + g_{\ell j} g_{jm}}{1 - g_{\ell j} g_{j\ell}}, & \ell, m \in J, \ell \neq m, g_{\ell j} g_{j\ell} < 1 \\ 0, & \text{otherwise} \end{cases}$$

3 Go to step 1.

Update algorithm

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:

$$J \to J/\{j\}$$

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

$$g_{\ell m} \to \begin{cases} \frac{g_{\ell m} + g_{\ell j} g_{jm}}{1 - g_{\ell j} g_{j\ell}}, & \ell, m \in J, \ell \neq m, g_{\ell j} g_{j\ell} < 1 \\ 0, & \text{otherwise} \end{cases}$$

3 Go to step 1.

Theorem

The initial levels α, the transition matrix **G** and the algorithm define a unique sequentially rejective test procedure that controls the FWER strongly at level α.

Proof idea:

- The graph and algorithm define weighted Bonferroni tests for each intersection hypothesis
- The algorithm defines a shortcut for the resulting consonant closed test, which does not depend on the rejection sequence
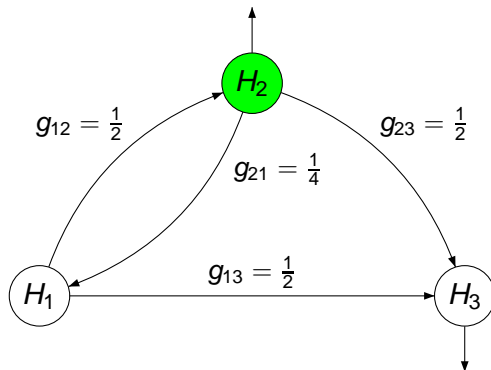
$g_{12} = \frac{1}{2}$

H_2

$g_{23} = \frac{1}{2}$

H_1

$g_{13} + g_{12}g_{23} = \frac{1}{2} + \frac{1}{2}\frac{1}{2} = \frac{3}{4}$

H_3

$g_{12} = \frac{1}{2}$

$g_{23} = \frac{1}{2}$

$g_{21} = \frac{1}{4}$

$g_{12} g_{21}$

H_2

H_1

H_3

$$\frac{g_{13} + g_{12} g_{23}}{1 - g_{12} g_{21}} = \frac{\frac{3}{4}}{1 - \frac{1}{2} \frac{1}{4}} = \frac{6}{7}$$

Generic Example

Example

- Two primary hypotheses H_1, H_2
  For example,
  - low/high dose for primary endpoint or non-inferiority claim
- Two secondary hypotheses H_3, H_4
  For example,
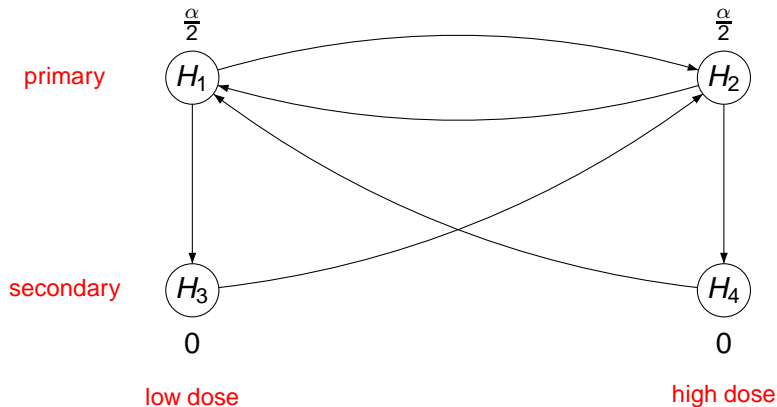 - low/high dose for secondary endpoint or superiority claim
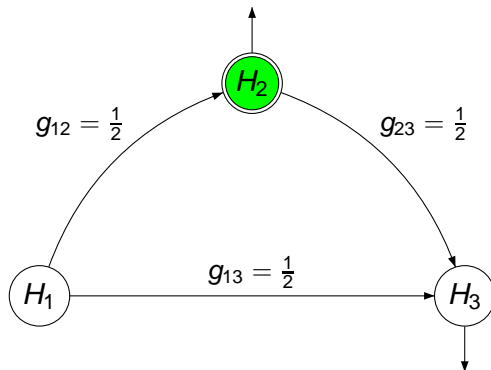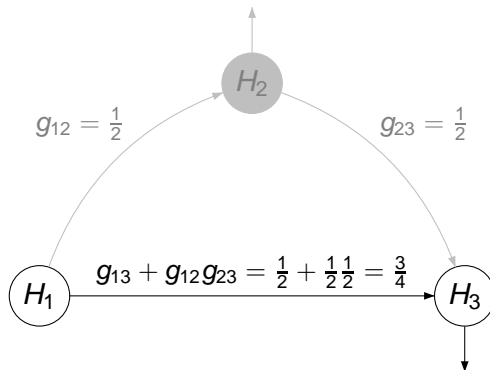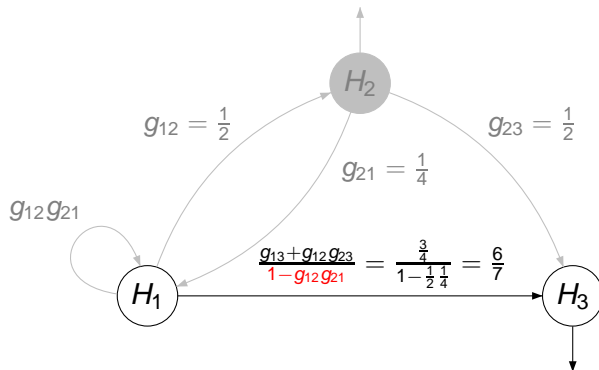
Successive test procedures for structured hypotheses

Example

- Two primary hypotheses H_1, H_2
  For example,
    - low/high dose for primary endpoint or non-inferiority claim
- Two secondary hypotheses H_3, H_4
  For example,
    - low/high dose for secondary endpoint or superiority claim

Proposed graphs

- ... are successive, control FWER, and display possible decision paths
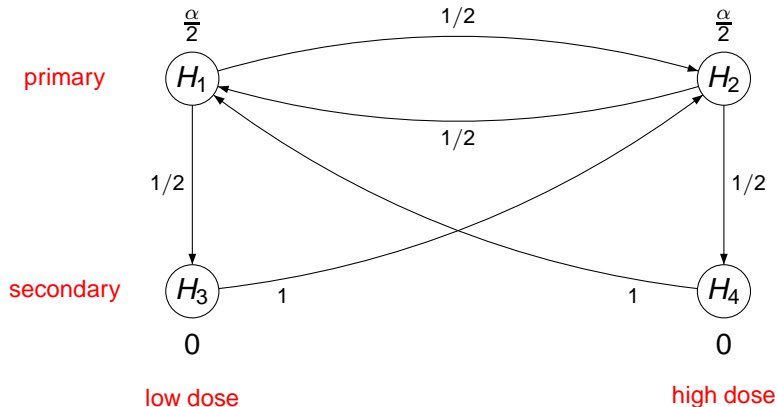- ... can be finetuned to reflect clinical considerations or treatment effect assumptions
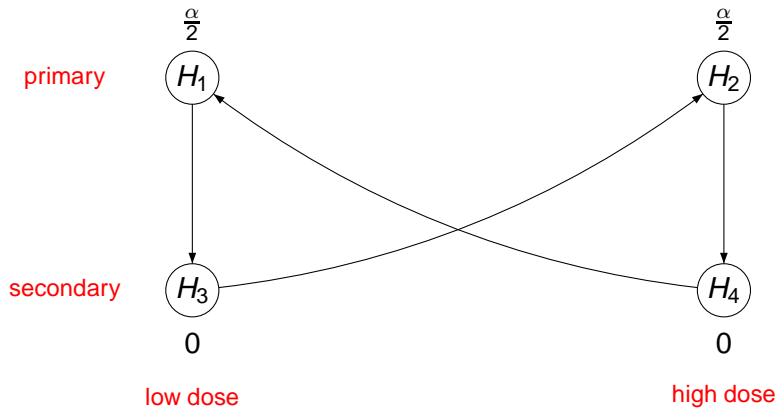
Example with $\alpha = 0.025$

primary

$p_1 = 0.01$

H_1

$\frac{\alpha}{2}$

1/2

H_2

$p_2 = 0.02$

$\frac{\alpha}{2}$

1/2

1/2

1/2

secondary

$p_3 = 0.07$

H_3

1

1

H_4

$p_4 = 0.001$

0

0

low dose

high dose

Example with $\alpha = 0.025$

primary

$p_1 = 0.01$

$p_2 = 0.02$

secondary

$p_3 = 0.07$

$p_4 = 0.001$

low dose

high dose

primary

$\frac{3\alpha}{4}$

H_2 $p_2 = 0.02$

1/3

2/3

1/2

secondary

$p_3 = 0.07$ H_3 1

$\frac{\alpha}{4}$

H_4 $p_4 = 0.001$

1/2   0

low dose

high dose
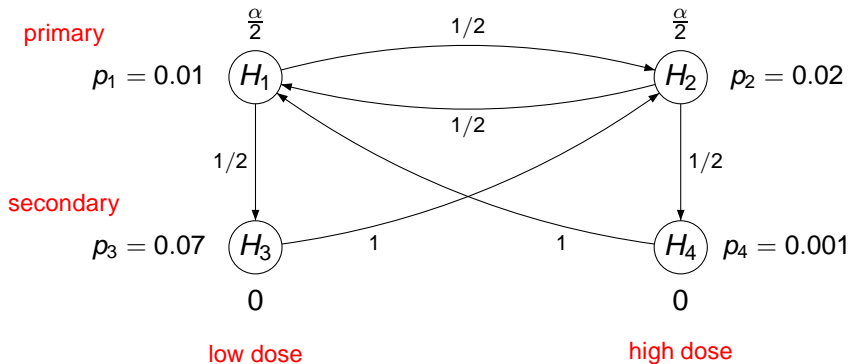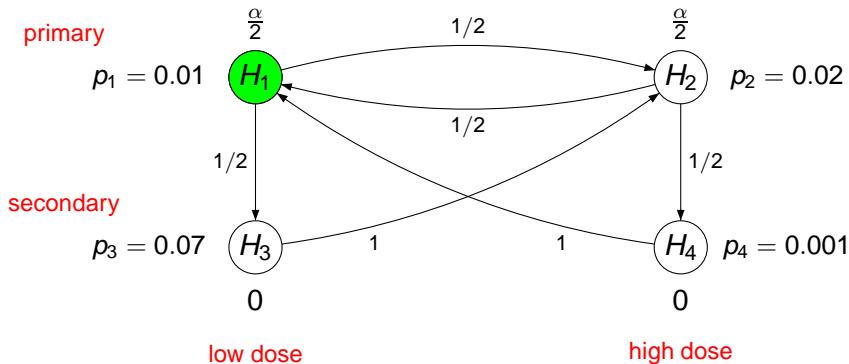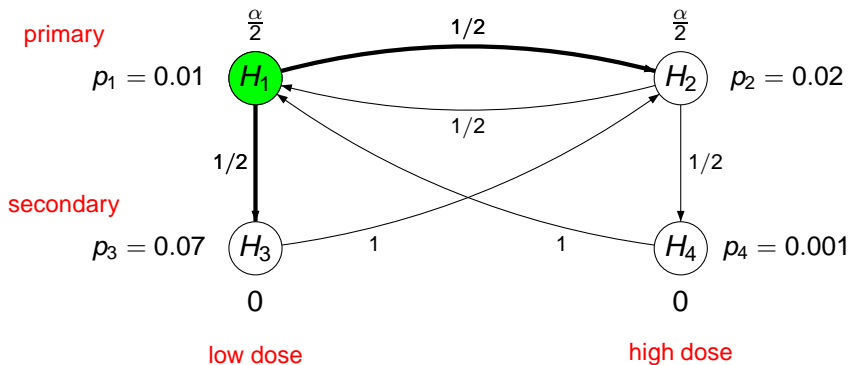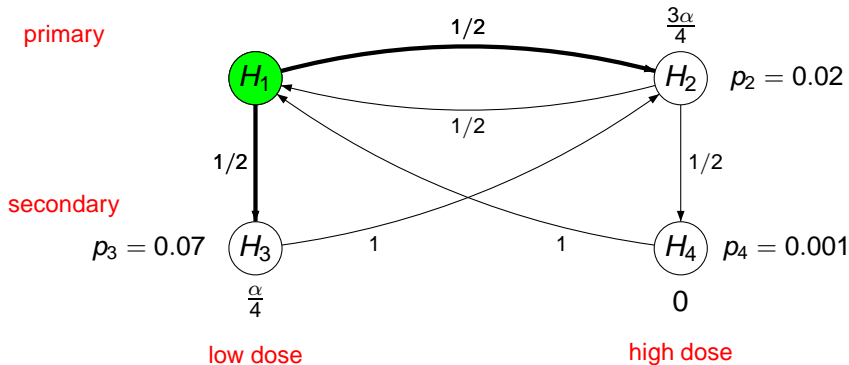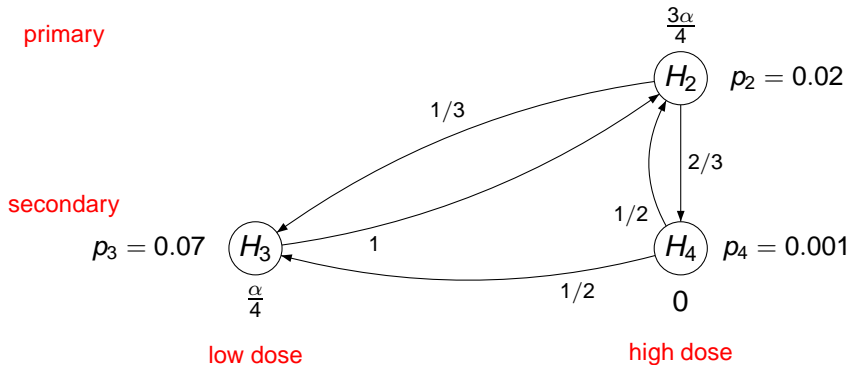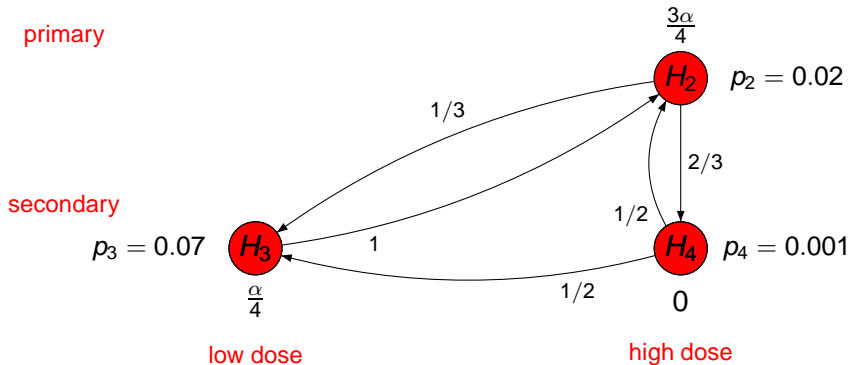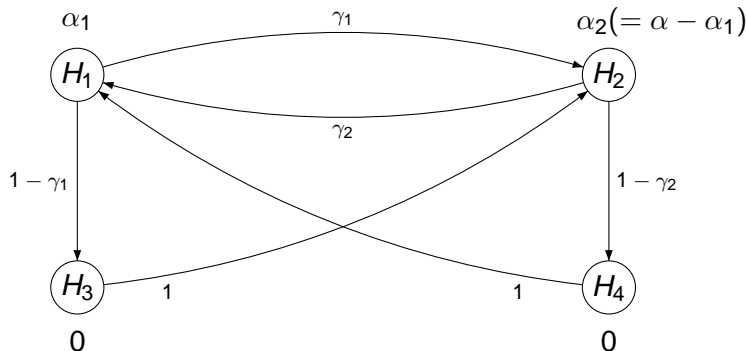
Resulting graph ...

needs to be finetuned with respect to α_1, γ_1, and γ_2, based on:

- further clinical considerations, or
- assumptions about effect sizes, correlations, etc.

Power considerations

Probability to reject at least one hypothesis, i.e. to identify at least one true effect

- depends only on the initial levels $\alpha_1, \ldots, \alpha_k$, and
- on the (unknown) true effect sizes and the correlations between the test statistics.
- For successive procedures only levels, effect sizes and correlations of primary hypotheses are relevant.

Probability to identify several true effects

- depends in addition on the edge weights of the graph.

Case Studies

Case study I

Late phase development of a new compound as an adjunctive therapy

Structured family of hypotheses

1. Four-armed trial comparing
   - Three dose levels of a new therapy adjunctive to standard-of-care
   - Placebo + standard-of-care as control
2. Two hierarchically ordered endpoints
   - Relapse rate and total medication score after 24 weeks

$\Rightarrow$ Six hypotheses H_{ij}

Dose $\quad i = 1$ (low), 2 (medium), 3 (high dose)

Endpoint $j = 1$ (relapse rate), 2 (total medication score)

(1) Relapse rate is more important than total medication score. Therefore,

- Primary hypotheses H_{11}, H_{21}, H_{31},
- Secondary hypotheses H_{12}, H_{22}, H_{32}.

(2) Primary hypotheses considered as equally important, but significance of adjacent doses (i.e. reject H_{11}, H_{21} or H_{21}, H_{31}) preferred over significance of non-adjacent significant doses (i.e. reject H_{11} and H_{31})

(3) Successiveness: A secondary hypothesis cannot be rejected without having rejected the associated parent primary hypothesis.

(1) Relapse rate is more important than total medication score. Therefore,

- Primary hypotheses H_{11}, H_{21}, H_{31}.
- Secondary hypotheses H_{12}, H_{22}, H_{32}.

(2) Primary hypotheses considered as equally important, but significance of adjacent doses (i.e. reject H_{11}, H_{21} or H_{21}, H_{31}) preferred over significance of non-adjacent significant doses (i.e. reject H_{11} and H_{31})

(3) Successiveness: A secondary hypothesis cannot be rejected without having rejected the associated parent primary hypothesis.

(1) Relapse rate is more important than total medication score. Therefore,
- Primary hypotheses H_{11}, H_{21}, H_{31},
- Secondary hypotheses H_{12}, H_{22}, H_{32}.

(2) Primary hypotheses considered as equally important, but significance of adjacent doses (i.e. reject H_{11}, H_{21} or H_{21}, H_{31}) preferred over significance of non-adjacent significant doses (i.e. reject H_{11} and H_{31})

(3) Successiveness: A secondary hypothesis cannot be rejected without having rejected the associated parent primary hypothesis.

(1) Relapse rate is more important than total medication score. Therefore,
  - Primary hypotheses H_{11}, H_{21}, H_{31},
  - Secondary hypotheses H_{12}, H_{22}, H_{32}.

(2) Primary hypotheses considered as equally important, but significance of adjacent doses (i.e. reject H_{11}, H_{21} or H_{21}, H_{31}) preferred over significance of non-adjacent significant doses (i.e. reject H_{11} and H_{31})

(3) Successiveness: A secondary hypothesis cannot be rejected without having rejected the associated parent primary hypothesis.

Clinical considerations

(1) Relapse rate is more important than total medication score. Therefore,
- Primary hypotheses H_{11}, H_{21}, H_{31},
- Secondary hypotheses H_{12}, H_{22}, H_{32}.

(2) Primary hypotheses considered as equally important, but significance of adjacent doses (i.e. reject H_{11}, H_{21} or H_{21}, H_{31}) preferred over significance of non-adjacent significant doses (i.e. reject H_{11} and H_{31})

(3) Successiveness: A secondary hypothesis cannot be rejected without having rejected the associated parent primary hypothesis.

(1)

relapse rate H_{11} H_{21} H_{31}

low dose   medium dose   high dose

(1)

	low dose	medium dose	high dose
relapse rate	H_{11}	H_{21}	H_{31}
total score	H_{12}	H_{22}	H_{32}

(1)

relapse
rate $\quad H_{11}$ $\quad\quad\quad\quad\quad H_{21}$ $\quad\quad\quad\quad\quad H_{31}$

total
score $\quad H_{12}$ $\quad\quad\quad\quad\quad H_{22}$ $\quad\quad\quad\quad\quad H_{32}$

$\quad\quad\quad$ 0 $\quad\quad\quad\quad\quad\quad$ 0 $\quad\quad\quad\quad\quad\quad$ 0

low dose $\quad\quad\quad$ medium dose $\quad\quad\quad$ high dose

Resulting multiple test procedure

(2)

relapse
rate

$\frac{\alpha}{3}$

(H_{11})

$\frac{\alpha}{3}$

(H_{21})

$\frac{\alpha}{3}$

(H_{31})

total
score

(H_{12})

0

(H_{22})

0

(H_{32})

0

low dose

medium dose

high dose

Resulting multiple test procedure

(2)

relapse rate: H_{11} $\xrightarrow{}$ H_{21} $\xleftarrow{}$ H_{31}, each with $\frac{\alpha}{3}$
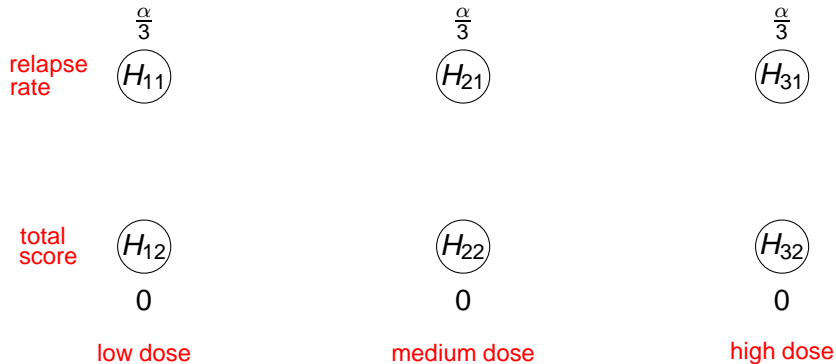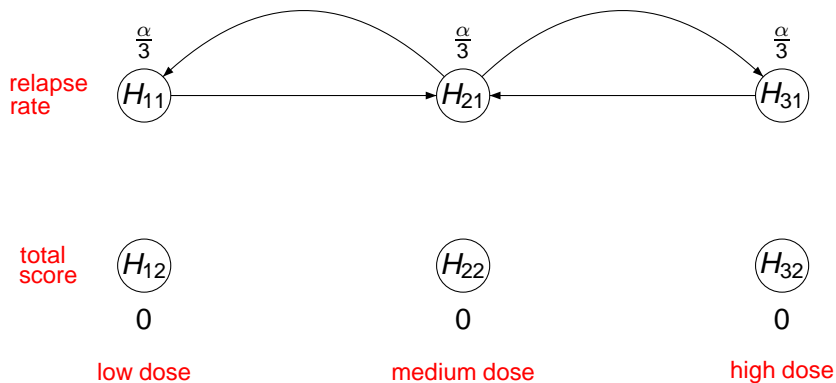
total score: H_{12} 0 (low dose), H_{22} 0 (medium dose), H_{32} 0 (high dose)

(3)
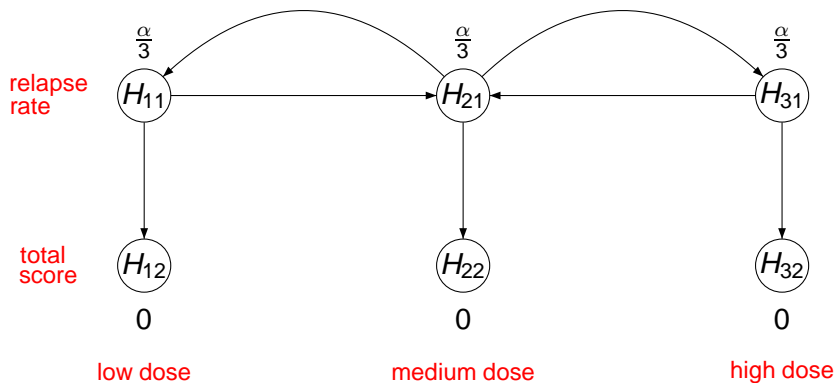
Resulting multiple test procedure
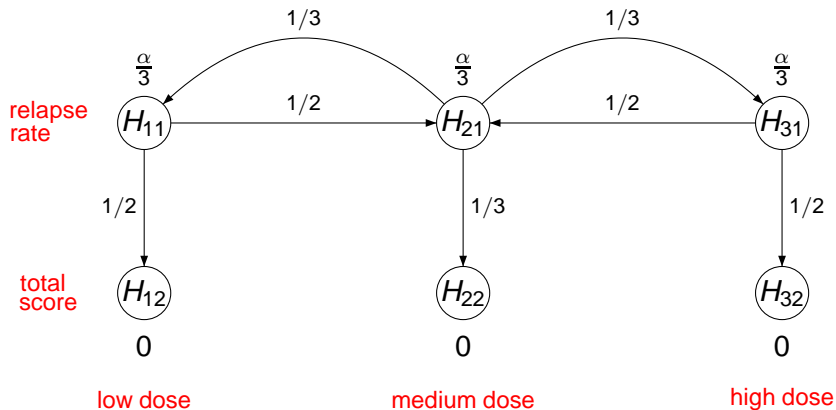
Resulting multiple test procedure

(2) + (3)
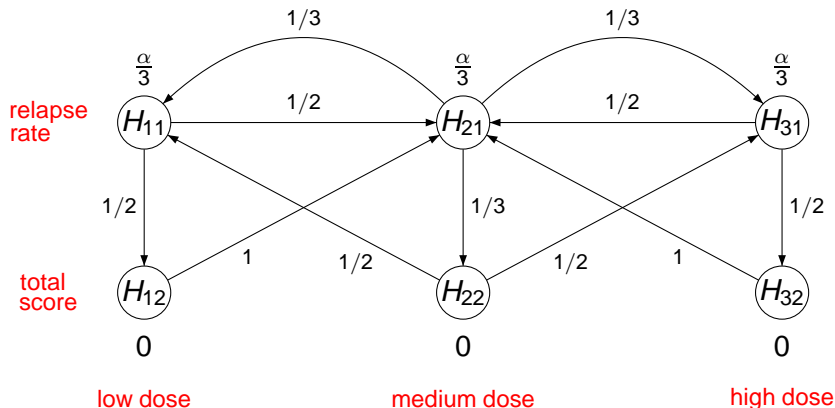
Easily implemented in SAS/IML

```
/* h: indicator whether a hypothesis is rejected (= 1) or not (= 0) (1 x n vector)
   a: initial significance level allocation (1 x n vector)
   w: weights for the edges (n x n matrix)
   p: observed p-values (1 x n vector) */

START mcp(h, a, w, p);
    n = NCOL(h);
    mata = a;

    crit = 0;
    DO UNTIL(crit = 1);
        test = (p < a);
        IF (ANY(test)) THEN DO;
            rej = MIN(LOC(test#(1:n)));
            h[rej] = 1;
            w1 = J(n, n, 0);
            DO i = 1 TO n;
                a[i] = a[i] + a[rej]*w[rej,i];
                IF (w[i,rej]*w[rej,i]<1) THEN DO j = 1 TO n;
                    w1[i,j] = (w[i,j] + w[i,rej]*w[rej,j])/(1 - w[i,rej]*w[rej,i]);
                END;
                w1[i,i] = 0;
            END;
            w = w1; w[rej,] = 0; w[,rej] = 0;
            a[rej]  = 0;
            mata = mata // a;
        END;
        ELSE crit = 1;
    END;

    PRINT h; PRINT (ROUND(mata, 0.0001)); PRINT (ROUND(w,0.01));
FINISH;
```

Example call

```
PROC IML;

START mcp(h, a, w, p);
    ...
FINISH;

/*** Numerical example ***/
h = {0 0 0 0 0 0};
a = {0.00833 0.00833 0.00833 0 0 0};
w = {0        0.5 0        0.5 0        0   ,
     0.3333 0   0.3333 0   0.3333 0   ,
     0        0.5 0        0   0        0.5,
     0        1   0        0   0        0   ,
     0.5    0   0.5    0   0        0   ,
     0        1   0        0   0        0   };
p = {0.1 0.008 0.005 0.15 0.04 0.006};

RUN mcp(h, a, w, p);
QUIT;
```
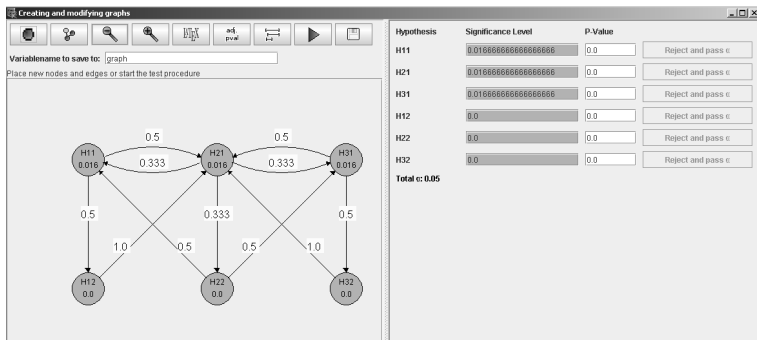
gMCP package in R

- Open source package available on CRAN at
  http://cran.r-project.org/web/packages/gMCP/
- Provides GUI within R through interface to JAVA

Case study II
Late phase development of a new cardiovascular drug

- Two treatments (A and B) compared with comparator (C)
- Superiority and non-inferiority tests for primary and multiple secondary endpoints.
- Three elementary hypotheses and two families of hypotheses:
  - H_1: superiority of A vs. C
  - H_2: non-inferiority of B vs. C
  - H_3: superiority of B vs. C
  - $\mathcal{H}_4$: multiple secondary variables for A vs. C
  - $\mathcal{H}_5$: multiple secondary variables for B vs. C

primary H_1 H_2

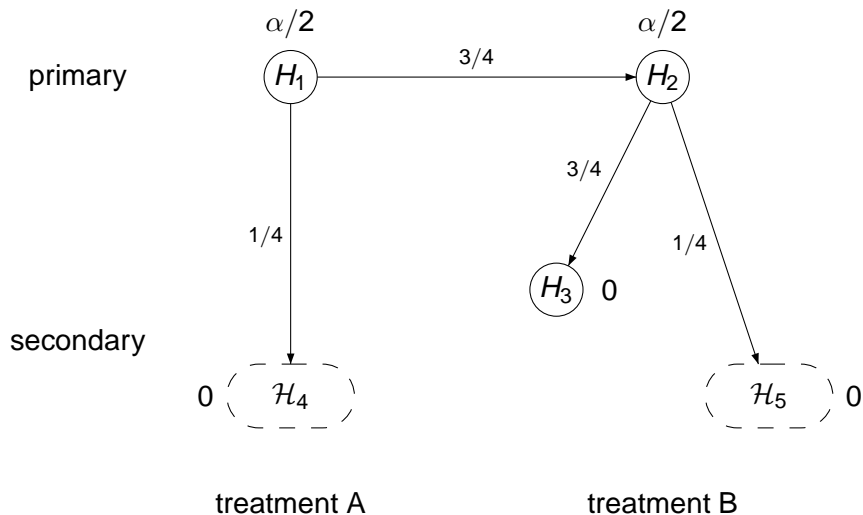treatment A   treatment B

Final multiple test procedure
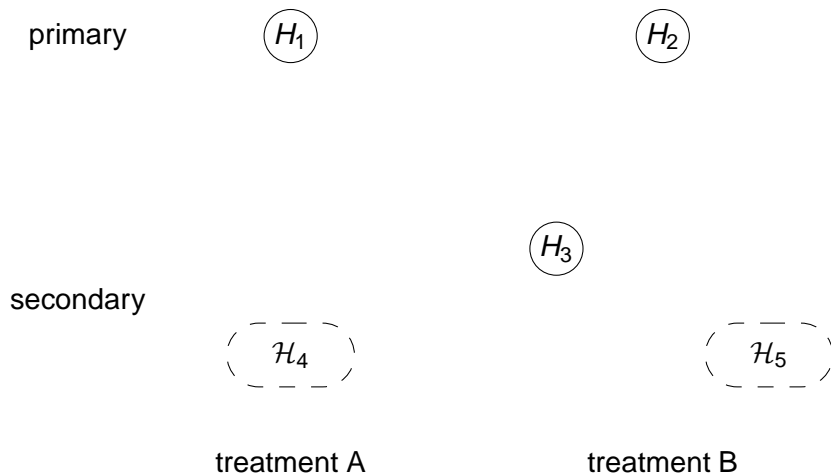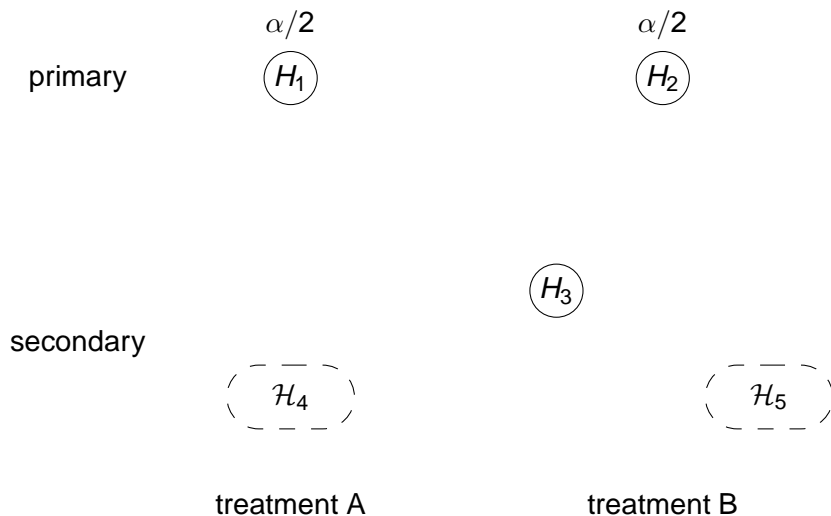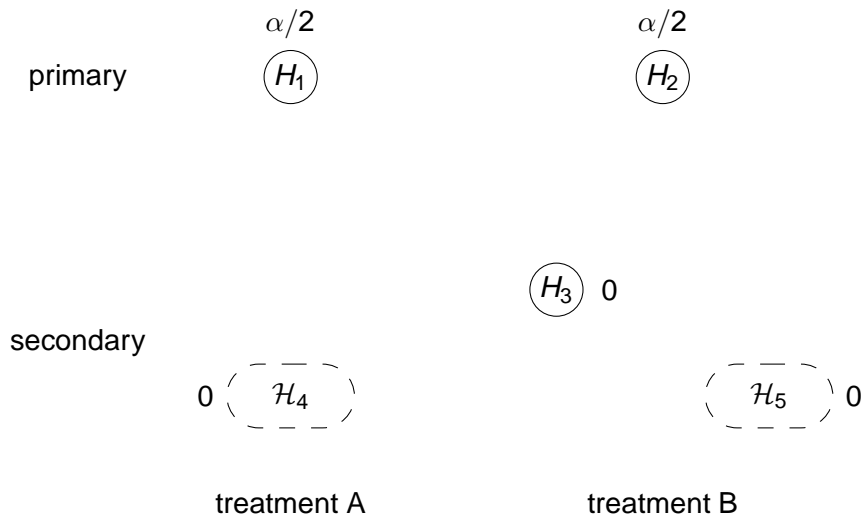
Final multiple test procedure
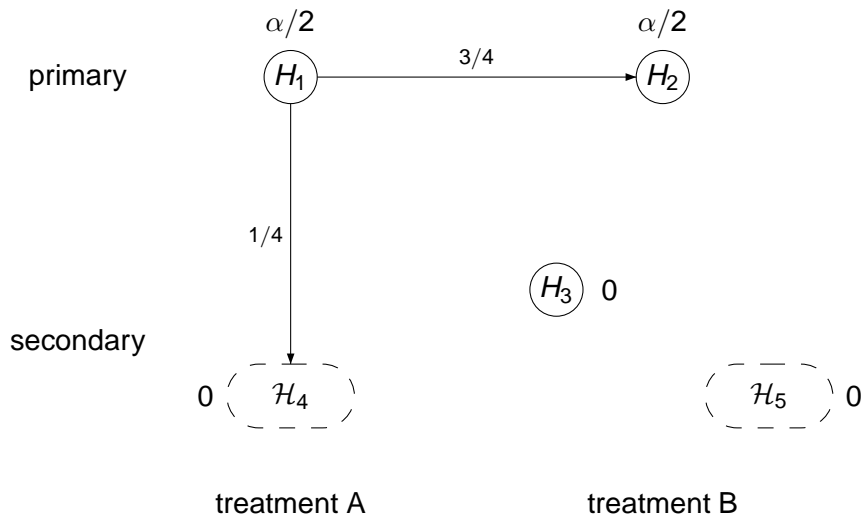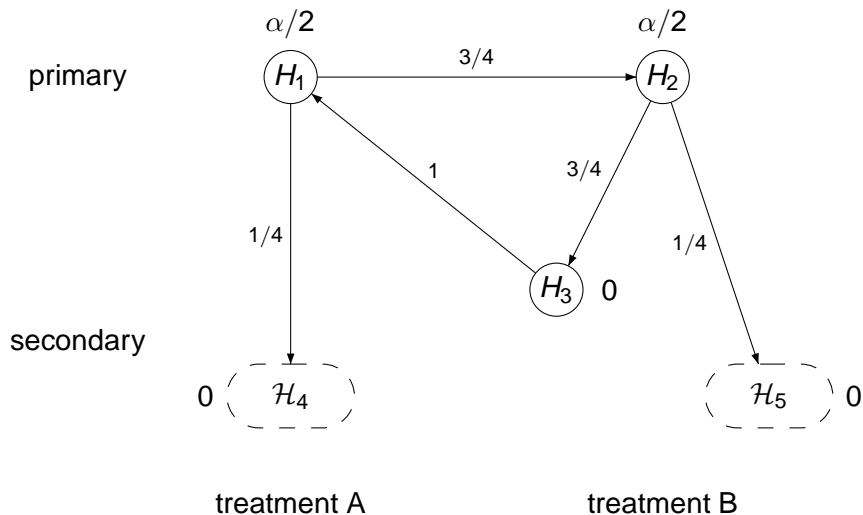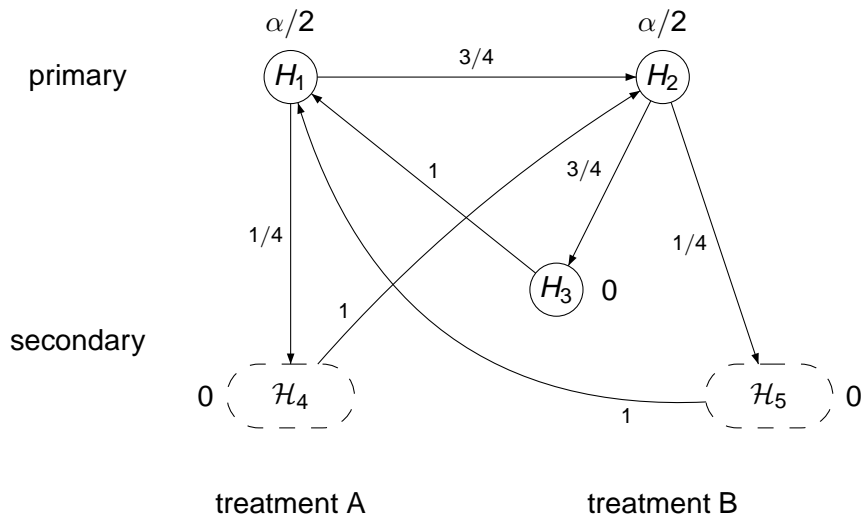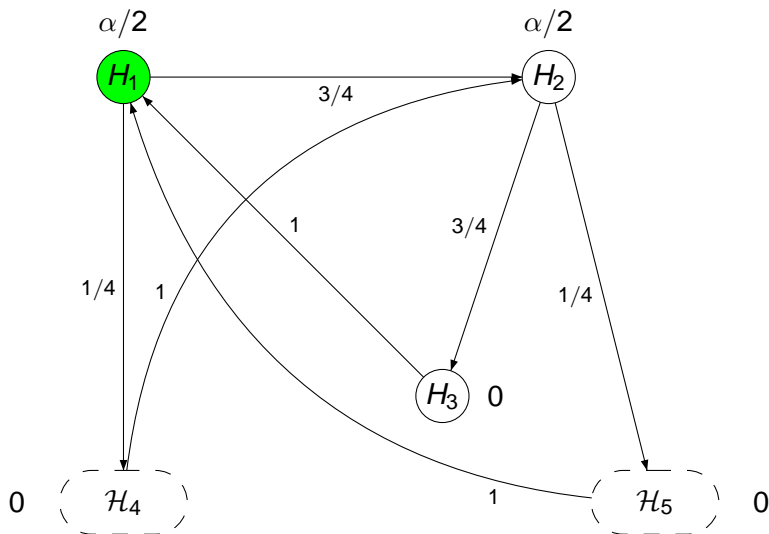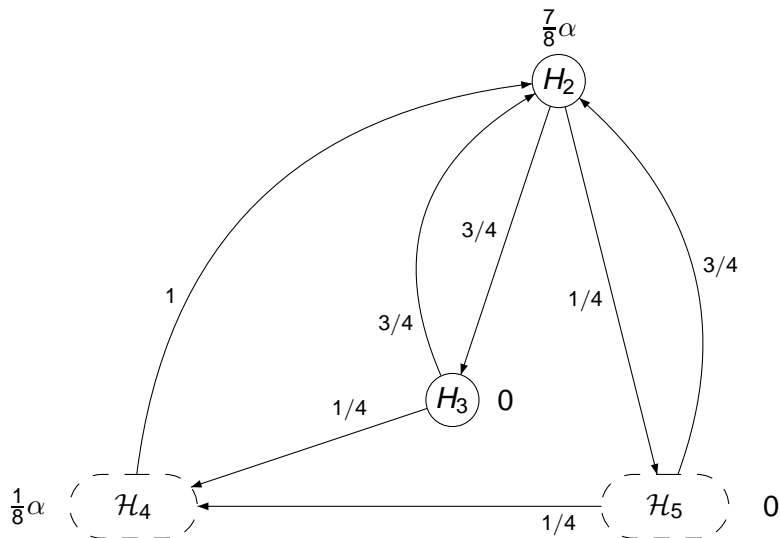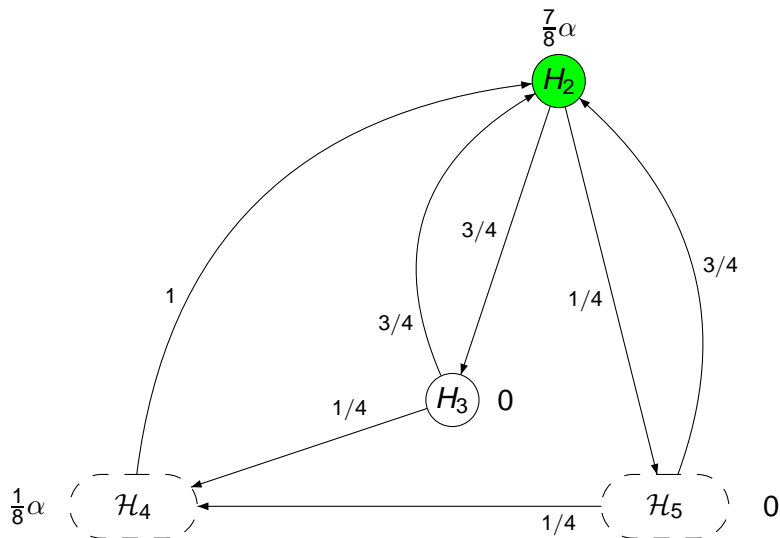
Final multiple test procedure

Final multiple test procedure

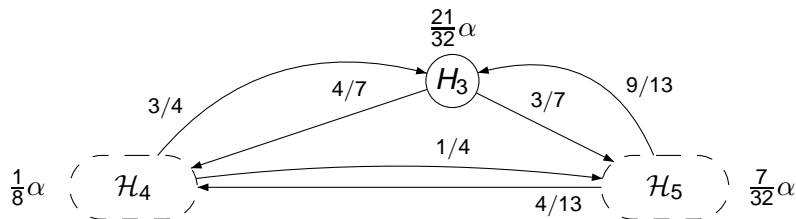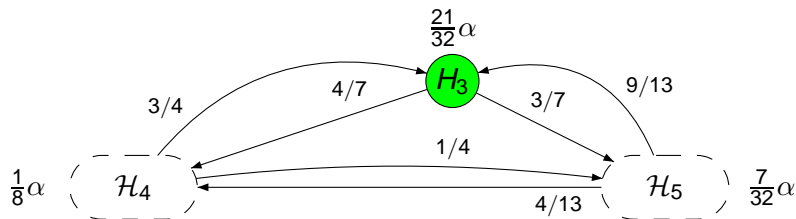# Final multiple test procedure

Example

Example

Example

Example

Example

Example
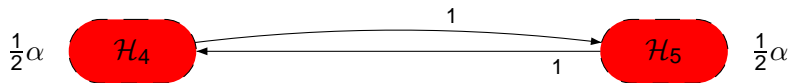
- Gold standard design comparing experimental drug (E) against placebo (P) and active comparator (C)
- Mixture of superiority and non-inferiority tests
- Two primary endpoints: pain, symptoms, resulting in six elementary hypotheses:

$H_{P,p}^{sup}$: superiority of E vs. P for pain

$H_{P,s}^{sup}$: superiority of E vs. P for symptoms

$H_{C,p}^{ni}$: non-inferiority of E vs. C for pain

$H_{C,s}^{ni}$: non-inferiority of E vs. C for symptoms

$H_{C,p}^{sup}$: superiority of E vs. C for pain

$H_{C,s}^{sup}$: superiority of E vs. C for symptoms

α 0   0   0   0   0

$H_{P,p}^{sup}$ —1→ $H_{C,p}^{ni}$ —1→ $H_{P,s}^{sup}$ —1→ $H_{C,s}^{ni}$ —1→ $H_{C,p}^{sup}$ —1→ $H_{C,s}^{sup}$
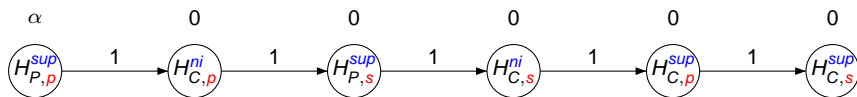
Next proposal: gatekeeping
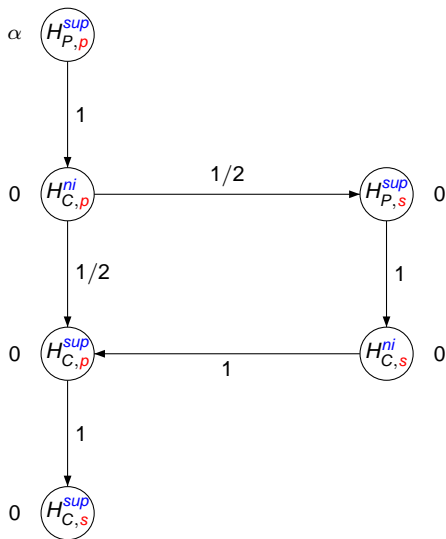
Next proposal: gatekeeping

Reasonable procedure?

pain   symptoms

pain

symptoms

sup of E vs. P

ni of E vs. C

sup of E vs. C

pain

symptoms

$H_{P,p}^{sup}$

$H_{P,s}^{sup}$

sup of E vs. P

$H_{C,p}^{ni}$

$H_{C,s}^{ni}$

ni of E vs. C

$H_{C,p}^{sup}$

$H_{C,s}^{sup}$

sup of E vs. C

pain

symptoms

α $\left(H_{P,p}^{sup}\right)$

$\left(H_{P,s}^{sup}\right)$ 0

sup of E vs. P

0 $\left(H_{C,p}^{ni}\right)$

$\left(H_{C,s}^{ni}\right)$ 0

ni of E vs. C

0 $\left(H_{C,p}^{sup}\right)$

$\left(H_{C,s}^{sup}\right)$ 0
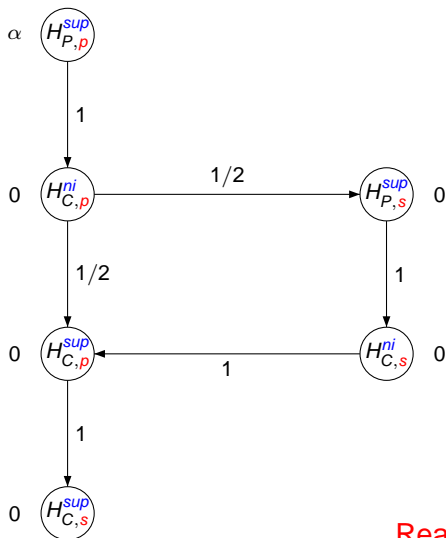
sup of E vs. C

Next proposal: gatekeeping – new display

pain   symptoms

α $H_{P,p}^{sup}$ $H_{P,s}^{sup}$ 0   sup of E vs. P

1   1/2

$H_{C,p}^{ni}$ 0 $H_{C,s}^{ni}$ 0   ni of E vs. C

1/2   1

0   $H_{C,p}^{sup}$ $H_{C,s}^{sup}$ 0   sup of E vs. C

1

Reject $H_{P,p}^{sup}$, $H_{C,p}^{ni}$, $H_{C,p}^{sup}$

pain　　　　　　　symptoms

$H_{P,s}^{sup}$　$\frac{\alpha}{2}$　　sup of E vs. P

1

$H_{C,s}^{ni}$　0　　ni of E vs. C

1

$H_{C,s}^{sup}$　$\frac{\alpha}{2}$　　sup of E vs. C

Reject $H_{P,p}^{sup}$, $H_{C,p}^{ni}$, $H_{C,p}^{sup}$

pain   **symptoms**

$H_{P,s}^{sup}$ $\frac{\alpha}{2}$ sup of E vs. P

1

$H_{C,s}^{ni}$ 0   ni of E vs. C

1
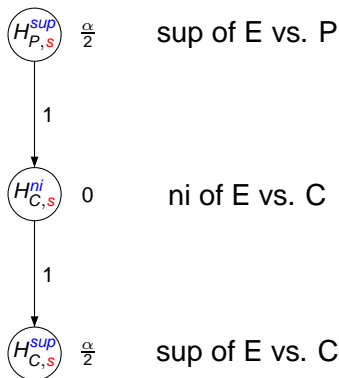
$H_{C,s}^{sup}$ $\frac{\alpha}{2}$ sup of E vs. C

Reject $H_{P,p}^{sup}$, $H_{C,p}^{ni}$, $H_{C,p}^{sup}$ $\Rightarrow$ Lack of succession property

Final proposal

pain

symptoms

α $\left(H_{P,p}^{sup}\right)$ $\left(H_{P,s}^{sup}\right)$ 0   sup of E vs. P

0   $\left(H_{C,p}^{ni}\right)$ $\left(H_{C,s}^{ni}\right)$ 0   ni of E vs. C

0   $\left(H_{C,p}^{sup}\right)$ $\left(H_{C,s}^{sup}\right)$ 0   sup of E vs. C

Final proposal
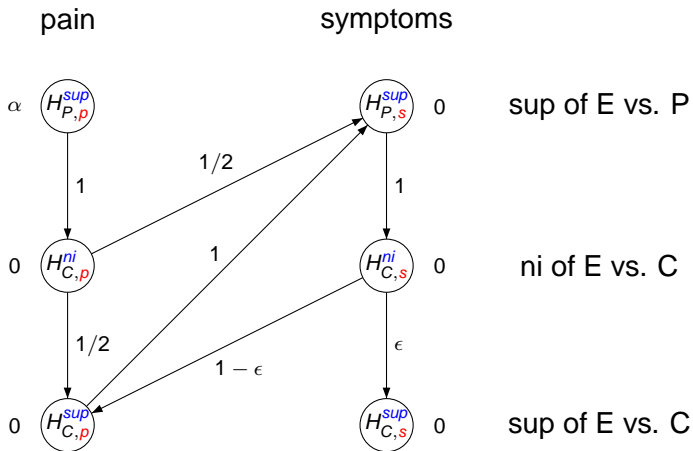
- Graph describing a test procedure for 4 EEG frequency bands and 5 sub-bands (Ferber et al., 2011)

G. Ferber et al. / Journal of Neuroscience Methods 201 (2011) 204–212
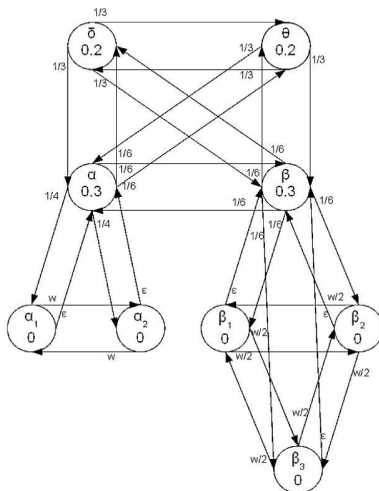
- Proposed graphical approach offers the possibility to
  - tailor advanced multiple test procedures to structured families of hypotheses,
  - visualize complex decision strategies in an efficient and easily communicable way, and
  - ensure strong FWER control.

- Approach covers many common gatekeeping procedures as special cases (Holm, fixed sequence, fallback, ...)

- Proposed graphical approach offers the possibility to
    - tailor advanced multiple test procedures to structured families of hypotheses,
    - visualize complex decision strategies in an efficient and easily communicable way, and
    - ensure strong FWER control.
- Approach covers many common gatekeeping procedures as special cases (Holm, fixed sequence, fallback, ...)

Summary

- Proposed graphical approach offers the possibility to
    - tailor advanced multiple test procedures to structured families of hypotheses,
    - visualize complex decision strategies in an efficient and easily communicable way, and
    - ensure strong FWER control.
- Approach covers many common gatekeeping procedures as special cases (Holm, fixed sequence, fallback, ...)

Summary

- Proposed graphical approach offers the possibility to
    - tailor advanced multiple test procedures to structured families of hypotheses,
    - visualize complex decision strategies in an efficient and easily communicable way, and
    - ensure strong FWER control.
- Approach covers many common gatekeeping procedures as special cases (Holm, fixed sequence, fallback, ...)

Summary

- Proposed graphical approach offers the possibility to
  - tailor advanced multiple test procedures to structured families of hypotheses,
  - visualize complex decision strategies in an efficient and easily communicable way, and
  - ensure strong FWER control.
- Approach covers many common gatekeeping procedures as special cases (Holm, fixed sequence, fallback, ...)

- Approach can be extended to address other problems:
  - Adjusted p-values and simultaneous confidence intervals available
  - Power and sample size considerations
  - Use of weighted and trimmed Simes tests
  - Weighted parametric test procedures to exploit correlation
  - Multiple testing in group sequential trials and adaptive designs
  - Convex combination of graphs and other geometric representation of multiple test procedures

- Approach can be extended to address other problems:
  - Adjusted p-values and simultaneous confidence intervals available
  - Power and sample size considerations
  - Use of weighted and trimmed Simes tests
  - Weighted parametric test procedures to exploit correlation
  - Multiple testing in group sequential trials and adaptive designs
  - Convex combination of graphs and other geometric representation of multiple test procedures

- Approach can be extended to address other problems:
  - Adjusted p-values and simultaneous confidence intervals available
  - Power and sample size considerations
  - Use of weighted and trimmed Simes tests
  - Weighted parametric test procedures to exploit correlation
  - Multiple testing in group sequential trials and adaptive designs
  - Convex combination of graphs and other geometric representation of multiple test procedures

Extensions

- Approach can be extended to address other problems:
  - Adjusted p-values and simultaneous confidence intervals available
  - Power and sample size considerations
  - Use of weighted and trimmed Simes tests
  - Weighted parametric test procedures to exploit correlation
  - Multiple testing in group sequential trials and adaptive designs
  - Convex combination of graphs and other geometric representation of multiple test procedures

- Approach can be extended to address other problems:
  - Adjusted p-values and simultaneous confidence intervals available
  - Power and sample size considerations
  - Use of weighted and trimmed Simes tests
  - Weighted parametric test procedures to exploit correlation
  - Multiple testing in group sequential trials and adaptive designs
  - Convex combination of graphs and other geometric representation of multiple test procedures

Extensions

- Approach can be extended to address other problems:
  - Adjusted p-values and simultaneous confidence intervals available
  - Power and sample size considerations
  - Use of weighted and trimmed Simes tests
  - Weighted parametric test procedures to exploit correlation
  - Multiple testing in group sequential trials and adaptive designs
  - Convex combination of graphs and other geometric representation of multiple test procedures

- Approach can be extended to address other problems:
  - Adjusted p-values and simultaneous confidence intervals available
  - Power and sample size considerations
  - Use of weighted and trimmed Simes tests
  - Weighted parametric test procedures to exploit correlation
  - Multiple testing in group sequential trials and adaptive designs
  - Convex combination of graphs and other geometric representation of multiple test procedures

Reference

- Bretz, Maurer, Maca (2014) Graphical approaches to multiple testing. To appear in: Young and Chen (eds.), *Clinical Trial Biostatistics and Biopharmaceutical Applications*, Taylor & Francis.