Alpha-recycling for the analyses of primary and secondary endpoints of clinical trials

Presenters:

Mohammad Huque, Ph.D.

FDA/CDER/OTS/ Office of Biostatistics

Sirisha Mushti, Ph.D.

Division of Biometrics V, Office of Biostatistics, OTS, CDER, FDA

BASS Conference, November 4, 2015, Rockville, Maryland

Acknowledgements

Lisa LaVange, Ph.D.
Office of Biostatistics, OTS, CDER, FDA

Mohamed Alosh, Ph.D.
Division of Biometrics III, Office of Biostatistics, OTS, CDER, FDA

Disclaimer

- ✓This presentation reflects the views of the presenters and should not be construed to represent FDA's views or policies.

Outline (Part I)

- Concepts and principles for testing multiple hypotheses of confirmatory clinical trials
- α-recycling concepts in testing multiple hypotheses
- Closed Testing Procedure (CTP)
  - How the CTP with the Weighted Bonferroni tests connects to the alpha-recycling and graphical methods?
- SR (sequentially rejective) graphical methods using:
  - Weighted Bonferroni tests
  - Weighted parametric tests for greater power
  - Simes tests for greater power
- Concluding Remarks (Part I)

Outline (Part II)

- Brief introductions on *B*-values and *Z*-scores used in GS (Group Sequential) test procedures and on alpha-spending functions (Ref: Proshan, Lan and Wittes; 2007)
- GS test procedures for testing multiple hypotheses
  - Methods based on the Bonferroni inequality
  - Method based on the CTP (Tang & Geller, 1999)
  - The case of testing 2 hypotheses
  - The general case of testing multiple hypotheses on using the graphical method (Maurer & Bretz, 2013)
- Example of a GS trial design for testing a primary and a secondary endpoint of a trial
- Concluding remarks (Part II)

Key References for Part I

- Huque MF, Dmitrienko A, and D'Agostino RB. Multiplicity issues in clinical trials with multiple objectives. *Statistics in Biopharmaceutical Research* 2013 (November)
- Bretz F (et al.) A graphical approach to sequentially rejective multiple test procedures. *Statistics in Medicine* 2009; 28: 586-604
- Bretz F (et al.) Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes or parametric tests. *Biometrical Journal* 2011; 53: 894-913

Key References for Part II

- Tang DI, Geller NL. Closed testing procedures for group sequential clinical trials with multiple endpoints. *Biometrics* 1999; 55: 1188-1192
- Maurer W and Bretz F. Multiple testing in group sequential trials using graphical approaches. *Statistics in Biopharmaceutical Research* 2013; 5(4): 311-320
- Yining Ye (et al.) A group sequential Holm procedure with multiple primary endpoints. *Statistics in Medicine* 2013; 32(7): 1112-1124
- Tamhane (et al.). Testing a primary and a secondary endpoint in a group sequential design. *Biometrics* 2010; 66: 1174-1184

2 books and 2 regulatory documents

- *Multiple Testing Problems in Pharmaceutical Statistics* - 2009

  Editors: A. Dmitrienko, A. C. Tamhane, and F. Bretz. Published by Chapman, and Hall/CRC Press, New York

  Chapter 1: Multiplicity Problems in Clinical Trials. A Regulatory Perspective (by Huque MF, and Röhmel J)

- *Multiple Comparison Using R* - 2010

  by Bretz, F., Hothorn, T., and Westfall, P; Published by CRC Press, New York

- CPMP/EWP/908/99. "Points to Consider on Multiplicity Issues in Clinical Trials,"

- FDA draft guidance on "multiple endpoints in clinical trials," 2015 (to be released)

References on GS methods

- *Statistical Monitoring of Clinical Trials*

  By Proschan, Lan and Wittes

  2007 print, by Springer (springer.com)

- *Group Sequential Methods*

  By Jennison & Turnbull

  Published in 2000 by Chapman &Hall/CRC, New York

Confirmatory clinical trials are generally designed with multiple objectives

- Primary objectives:
  - If the trial wins on one or more primary objectives, then one can characterize clinically relevant benefits of the study treatment
  - These objectives are defined in terms of the so called "primary endpoints" (PEs)
- Secondary objectives:
  - These are for describing additional clinically pertinent benefits of the study treatment. The secondary objectives are defined in terms of the so called "secondary endpoints" (SEs).
- Other objectives (e.g., tertiary, supportive, and exploratory)

Primary vs. secondary endpoints:

- They differ in concept and purpose
  - ✓ Efficacy of a treatment is derived on demonstrating clinically meaningful and statistically significant benefits of the study treatment in one or more primary endpoints satisfying a pre-defined clinical win scenario.
  - ✓ In general, SEs alone are not suitable for this special purpose.
  - ✓ SEs are generally used for establishing treatment benefits in addition to those already established by one or more PEs
- Reference:

  O'Neill RT. Secondary endpoints cannot be validly analyzed if the primary endpoint does not demonstrate clear statistical significance. *Controlled Clinical Trials* 1997; **18**: 550-556.

Multiplicity in clinical trials

- Multiplicity in a clinical trial arises when the trial design allows to win for efficacy or safety in multiple ways
  - ✓ Causes the Type I error rate to inflate requiring statistical adjustments for its control
  - ✓ There are useful statistical approaches to handle this
- **Example**: Consider a clinical trial that is designed to compare a new treatment to control for showing that this new treatment is superior to control in any of the three specified primary endpoints.

Multiplicity in clinical trials (cont'd)

- **Example**: Consider a drug trial for Alzheimer's disease that compares a new drug to control on two PEs:
  - ADAS-Cog (cognition scale) and
  - CIBIC plus (clinical global scale).
- Clinical win criterion: Statistical test for the treatment effect needs to be statistically significant at the 0.025 level (by 1-sided test) for each specified endpoint.
- Is there a multiplicity issue here?

Multiplicity in clinical trials (cont'd)

- **Example**: Consider a drug trial for epilepsy that compares a new treatment to control on 3 PEs:

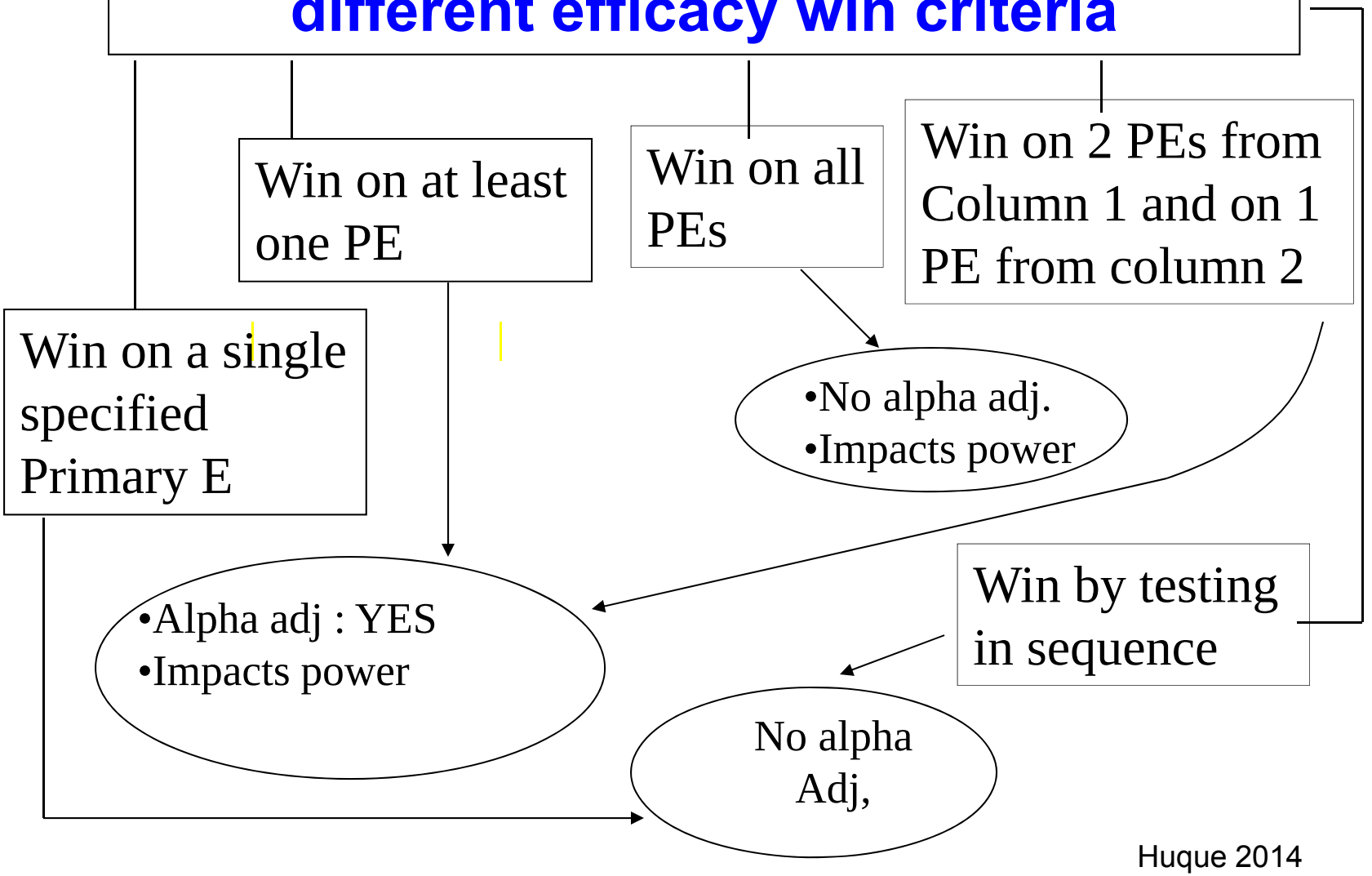
  A= seizure rate

  B= drop attack rate

  C= seizure severity

- Clinical win criterion: Show benefit of the study treatment either for A or for both B and C (Dmitrienko, D'Agostino, and Huque; 2013)
- Is there a multiplicity issue here?

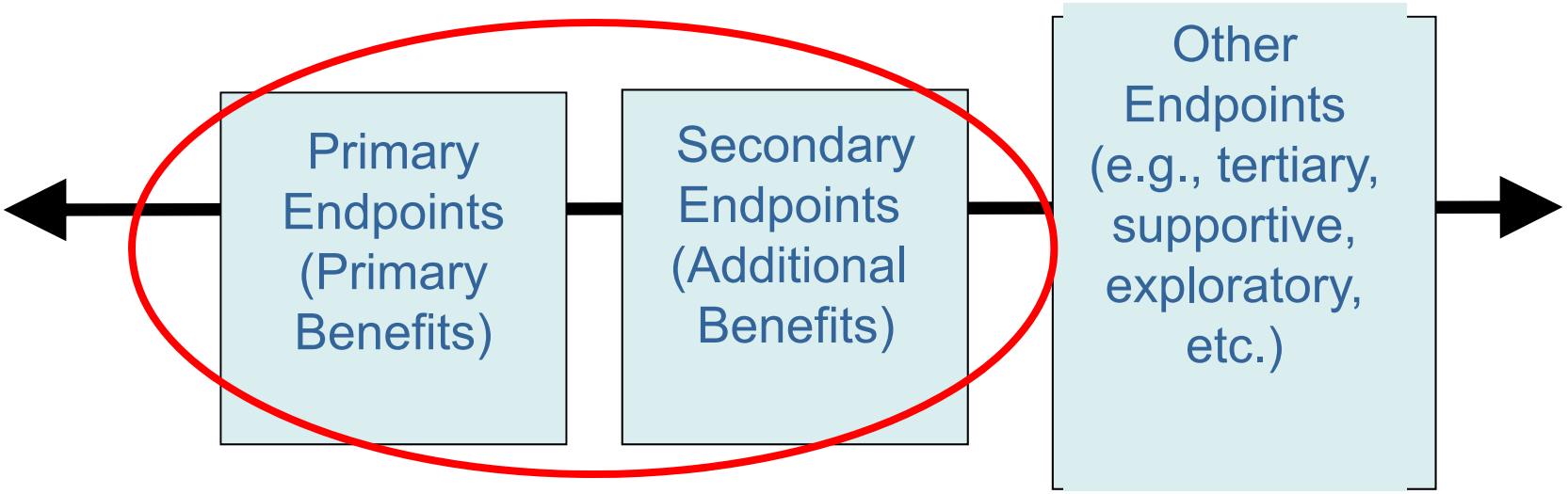
Clinical trial designs often come with different efficacy win criteria



Hypotheses in confirmatory trials usually follow a hierarchical structure

- PE hypotheses are considered more important
  - SE hypotheses are usually tested for statistical significance after there is a favorable clinically meaningful and statistically significant result involving one or more PEs
  - Statistical approaches for clinical trials are therefore tailored to this hierarchical structure, normally optimizing the power for testing the PE hypotheses
- For confirmatory trials, the use of standard methods such as Bonferroni, Holm, Hochberg, Dunnett t-tests, etc., on ignoring such hierarchical structures of test hypotheses, are generally considered inefficient

Continuum for the overall Type I error rate control



Overall error rate should not exceed a pre-specified level α

Statistical methods for confirmatory trials

- Statistical methods used are those that control overall Type I error rate (FWER) in the strong sense across both the primary and secondary families of hypotheses, so that conclusions of treatment benefits can be made at the individual hypothesis level
- Statistical methods that control FWER only in the weak sense is generally not considered.
- For confirmatory trials, hardly ever one is interested in the whether all hypotheses are jointly true or not.

Reference: Hochberg and Tamhane (1987)
for weak vs. strong FWER control definitions

Consequence of analyzing each secondary endpoint at the 0.05 level

- A practice has been to analyze a number of secondary endpoints each at the 0.05 level after successful results on one or more primary endpoints.
- This practice can have high inflation of the FWER (except for a very special case when these secondary endpoints are tested by the fixed sequence method after successful results on all specified primary endpoints).

Example 1

- Consider treatment-to-control comparisons in a trial on 4 endpoints (Dmitrienko, D'Agostino, and Huque; 2013):

  A is primary

  B, C and D are secondary

- Test strategy:

  Test for A at level 0.05

  If the test for A is significant, then test for B, C, and D each at level 0.05

Example 1 (cont'd)

- Suppose that the global null hypothesis is true, i.e., there is no treatment effects for any endpoint:
  - Then the probability of falsely concluding treatment effect in any endpoint = 0.05. That is FWER = 0.05.
  - Why? Because, tests for endpoints B, C, and D occur only after the test for endpoint A is significant at level 0.05. This renders the size of error rate for secondary endpoints not to exceed 0.05
- Why is then a problem?

Example 1 (cont'd)

- The previous calculation focused only on one null hypothesis configuration of true and false null hypotheses
- **Doing this can lead to a substantial underreporting of true error rate!!!**
- For example, consider the configuration:
  - The null hypothesis for A is false but those for B, C, and D are true
  - Then the error rate for the test strategy can be as high as $1 - (1 - 0.05)^3 = 0.142$ (on assuming tests are independent)
- If 5 secondary endpoints then FWER = 0.226

Issue of alpha for the secondary endpoint family

Should the secondary endpoint family be always analyzed at the full alpha level (e.g., at 0.05) after the trial is successful on one or more specified primary endpoints?

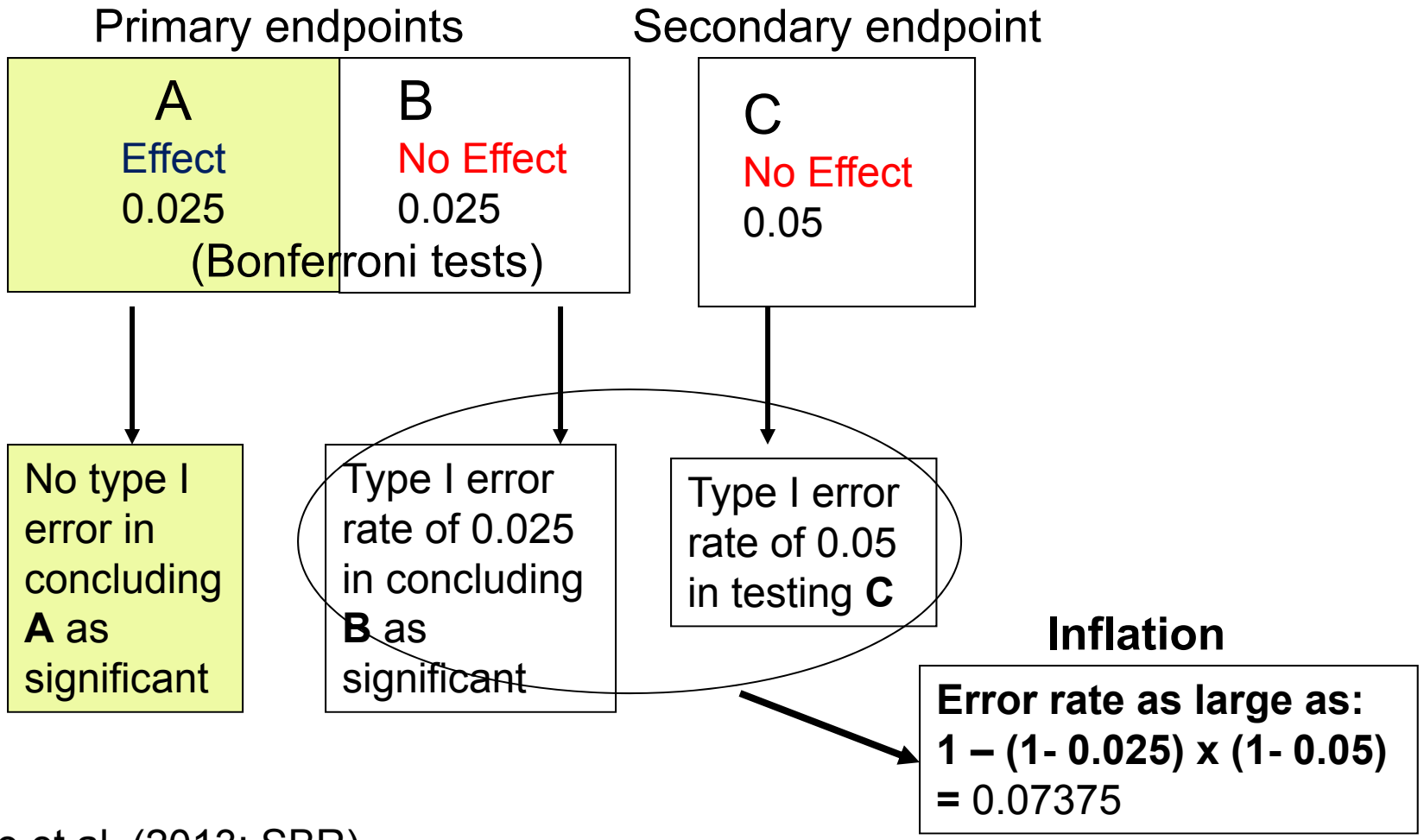
Issue of alpha for the secondary endpoint family (cont'd)

- If the trial has a single PE and several SEs, and if the trial is successful on that PE then full alpha is available for the secondary endpoint family.
- If the trial has two or more PEs and the trial is successful on all specified PEs then also full alpha is available for the secondary endpoint family. (follows from the gate-keeping test strategy)
- What about the situation when the trial is successful on some but not on all specified primary endpoints? Can the secondary endpoint family be assigned full alpha?

Example 2

- Consider a 2-arm trial designed to compare a treatment to control on two PEs (A and B) and on single secondary endpoint C
- Suppose that the Bonferroni method is applied for testing for A and B with each test at level 0.025, on splitting the trial alpha of 0.05
- Suppose that at the conclusion of the trial the observed treatment effect p-values are: $p_A < 0.001$ and $p_B = 0.20$.
- **Question**: Should there be full alpha of 0.05 available for this case for testing for the secondary endpoint C?

Example 2: Test PEs A and B, each at level 0.025, if win in one of them, then tests the secondary endpoint C at level 0.05



Huque et al. (2013; SBR)

Confirmatory trial results based on exploratory analyses are considered inconclusive. Why?

- Any conclusion of favorable result has a very high probability of false positive error
- Besides this high probability, results include serious bias components
- Interpretation of p-value is problematic

Examples of serious bias components in exploratory analyses

1) There is always a desire to report favorable result (conflict of interest bias)

2) Biological plausibility in favor of treatment is usually suggested after the result is seen and not before

3) With many analyses, each producing an estimate with variability, one pics the one which is most favorable. This produces random high bias which increases with the number of analyses and increase in variability

Exploratory analyses (cont'd)

- Thus, exploratory analyses are usually hypotheses generating exercises.
- Putting their results in the drug labels, in medical journals, and other publications, somehow to be used for promotional purposes, is problematic.
- Such a practice can have substantial misleading consequences.

A paradox noted by a statistician

- A statistician visited a hospital nursery about 45 years ago. The nursery was the central gathering place for the hospital's newborns in those days. He was surprised to observe that there were 20 babies of 1 sex and only two of the other.
- He computed a *P* value for the likelihood that an imbalance this extreme would have occurred by chance if indeed there were an equal sex distribution in the population at birth
- The 2-sided *P* value came out to be 0.0001 which he saw to be correct.
- Then what could explain this paradox?

Reference: *Clinical trials*: discerning hype from substance (*Ann of Intern Med* 2010; 153: 400-406; by Thomas Fleming)

Statistician's explanation of the paradox

- **"I did not walk into the hospital with the intention to gather prospective data to assess and report on this hypothesis. Rather, the data generated the hypothesis."**

Two key statistical approaches for the analyses of the PE and SE hypotheses of clinical trials

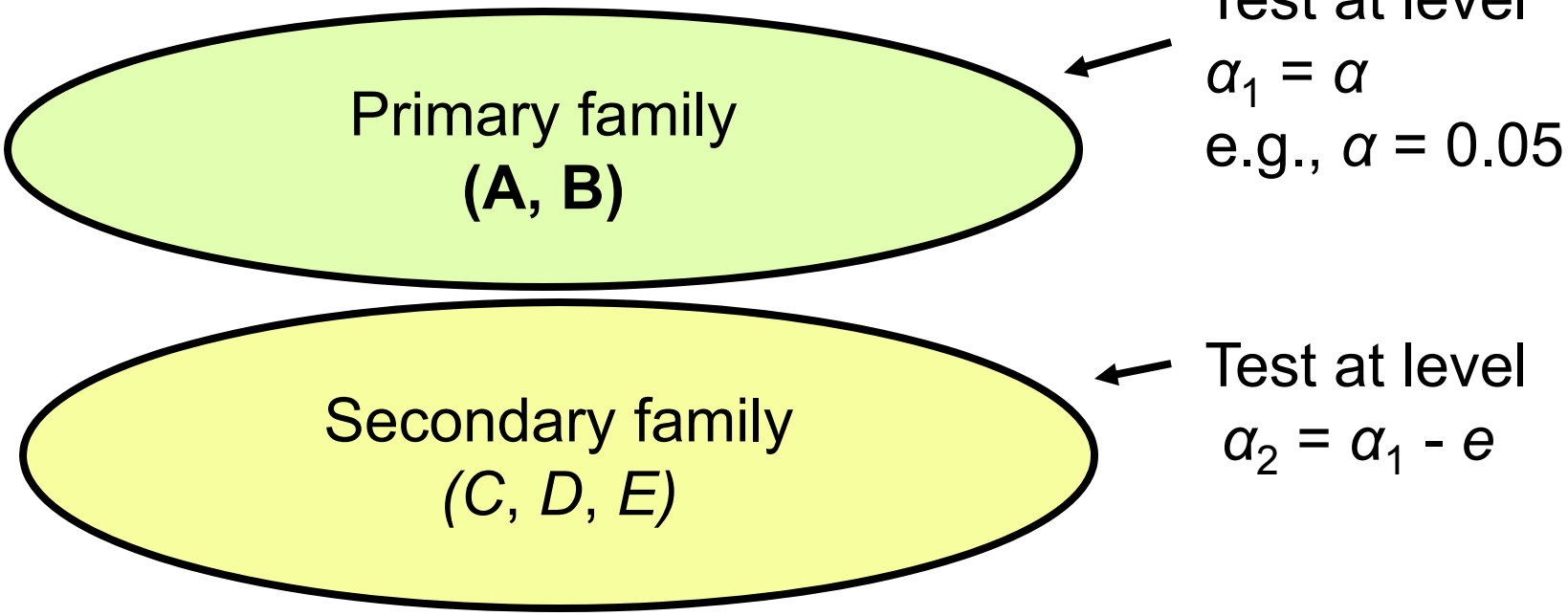
- Gatekeeping methods:
  - Dmitrienko A, D'Agostino RB, and Huque MF. Key multiplicity issues in clinical drug development, *Statistics in Medicine* 2013; 32: 1079 –1111
  - Huque MF, Dmitrienko A, and D'Agostino RB. Multiplicity issues in clinical trials with multiple objectives. *Statistics in Biopharmaceutical Research* 2013 (November)
- Graphical Methods:
  - Bretz F (et al.) A graphical approach to sequentially rejective multiple test procedures. *Statistics in Medicine* 2009; 28: 586-604
  - Bretz F (et al.) Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes or parametric tests. *Biometrical Journal* 2011; 53: 894-913
  - Maurer W and Bretz F. Multiple testing in group sequential trials using graphical approaches. *Statistics in Biopharmaceutical Research* 2013; 5(4): 311-320

Gatekeeping test strategy

- Useful for testing primary and secondary families of endpoints or hypotheses
- The usual strategy is to test all endpoints in the primary family by a method such as Bonferroni and proceed to the secondary family of endpoints only if there has been statistical success in the primary family.
- **This allows all of the trial alpha to be used for the primary family. Thus, maximizing the study power for those critical endpoints.**

Gatekeeping approach

- Consider two families of endpoints (or hypotheses), one primary and the other secondary

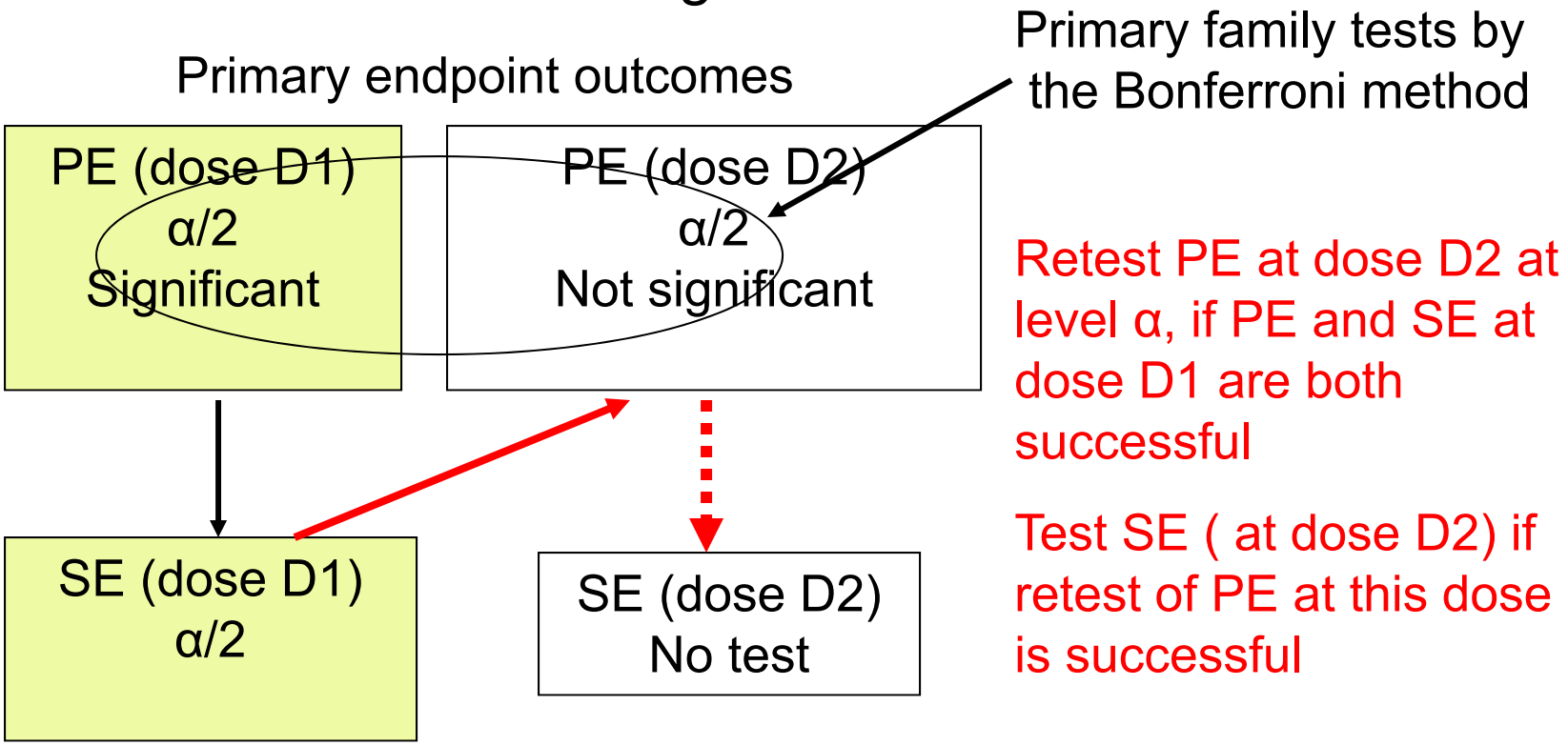


"e" depends on how many endpoints in the primary family are successful. If all endpoints are successful in this family then e = 0.

Gatekeeping approach w. re-testing

Consider 2 endpoints (PE, SE) and 2 dose levels D1 and D2

Logical restriction: SE at dose D1 or D2 cannot be tested unless PE at that dose is significant

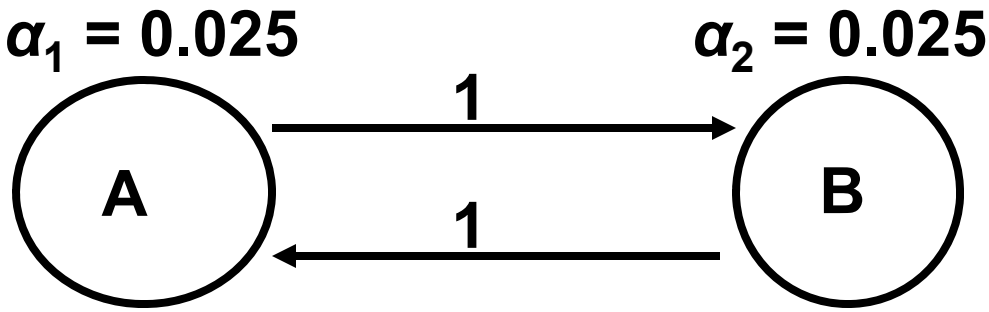


Retest PE at dose D2 at level α, if PE and SE at dose D1 are both successful

Test SE (at dose D2) if retest of PE at this dose is successful

Graphical approach is based on the concepts of "alpha saved" and "alpha lost"

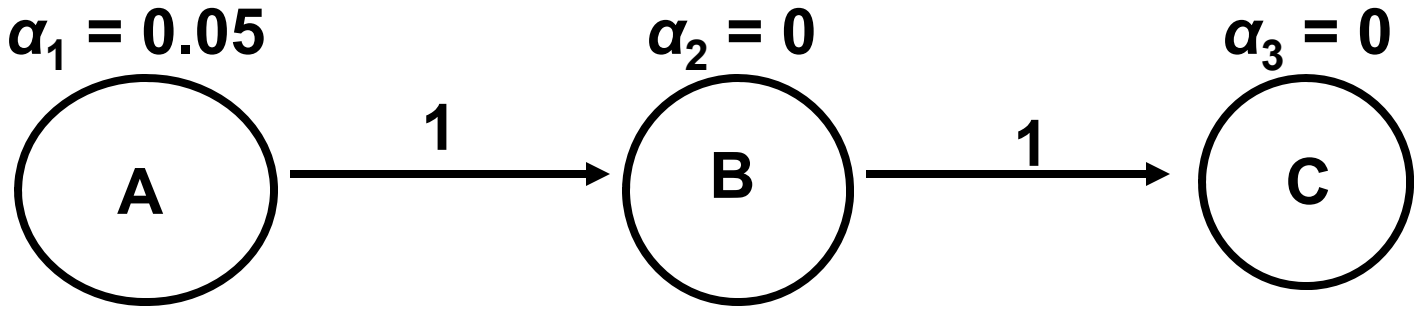
- If an endpoint (or hypothesis) is tested at a level alpha (e.g., alpha = 0.025) and the *p*-value is significant at that level then that alpha of 0.025 is "saved" and can be accumulated to test a second prospectively specified endpoint (or hypothesis)



Thus, if A is successful, then alpha at B is 0.025 +1*0.025 = 0.05

(This graph is the graphical representation of the Holm's test for testing two hypotheses)

Graphical representation of the fixed sequence (FS) method

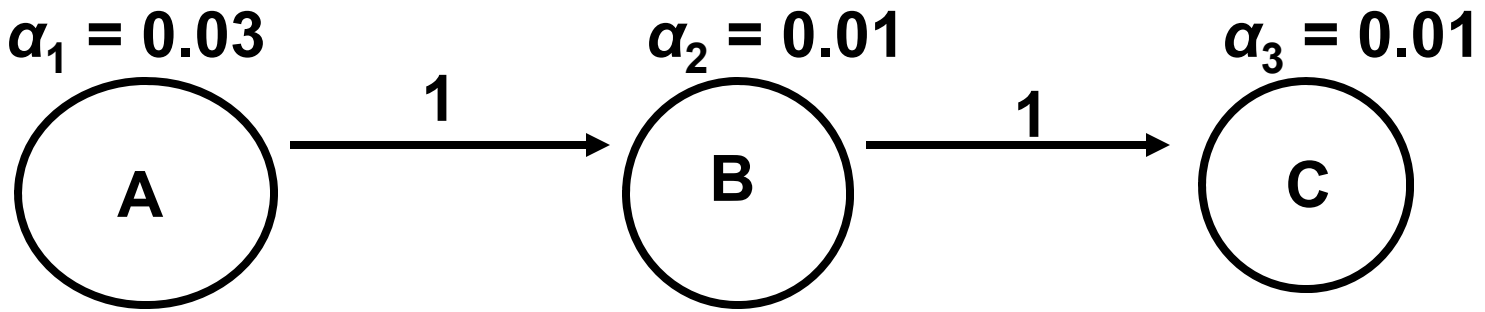


If A is successful, alpha for B becomes 0 +1*0.05 = 0.05. Then, if B is successful alpha for C is 0.05. But, if anytime, a test is not significant there is no further test

Drawback of the FS method

- If a hypothesis in the sequence is not rejected then a statistical conclusion cannot be made for the subsequent hypotheses, even if they have extremely small *p*-values.
  - Suppose, for example, that in a study the *p*-value for the first hypothesis test in the sequence is $p = 0.250$, and the *p*-value for the second hypothesis test is $p = 0.00001$.
  - Despite the apparent "strong" finding for the second hypothesis, no formal favorable statistical conclusion can be reached for this hypothesis.

Fallback method and its graphical representation



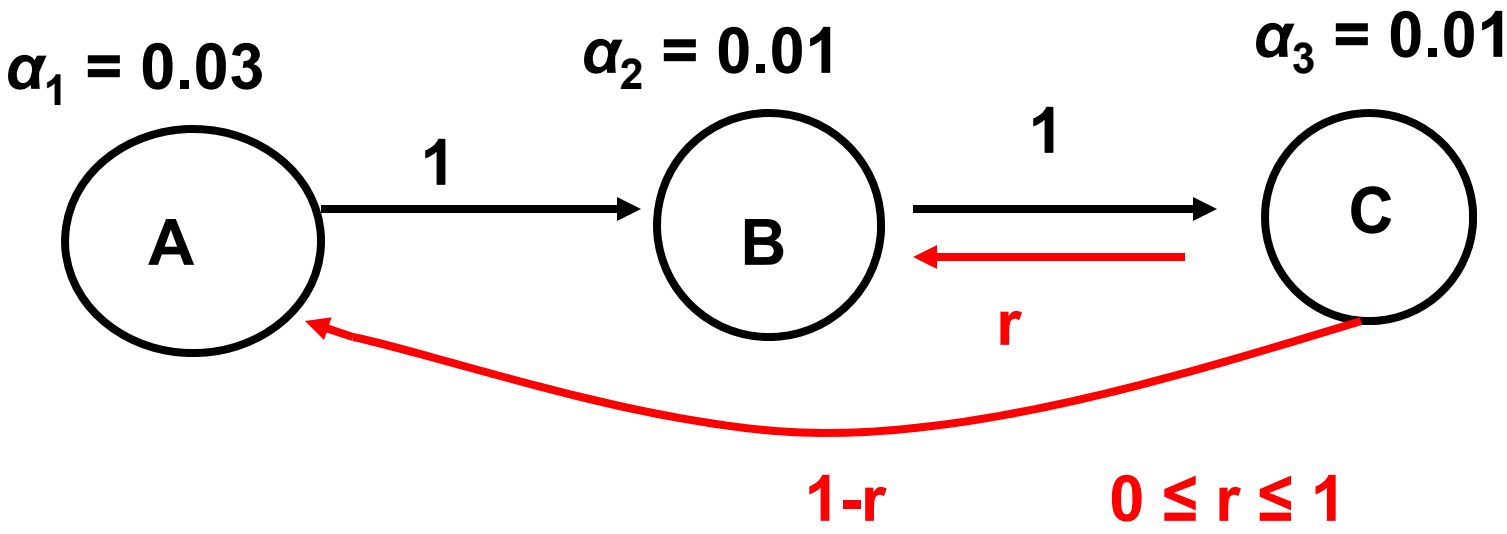
If A is successful, alpha for B becomes 0.01 +1*0.03 = 0.04, and if B is also successful, then test for C is at level 0.05 (This test strategy is known as the fallback method)

Reference: Brian Wiens (2003)

Extension of the fallback method

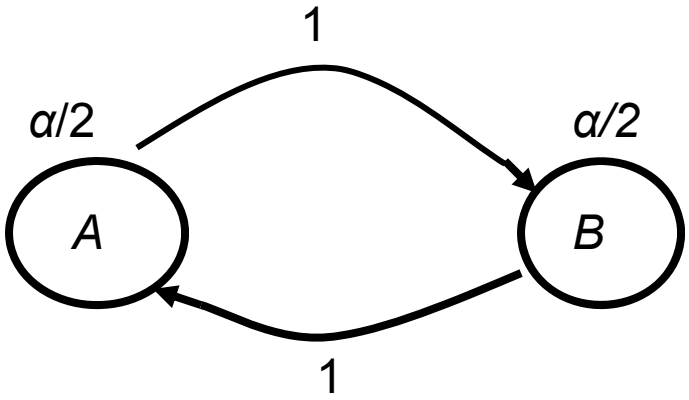
(Bretz et al.; 2009)

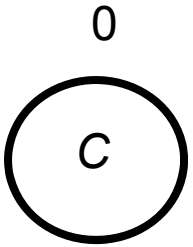
Consider the situation: A and B both fail but C is successful



Then A and B can be retested at slightly higher levels

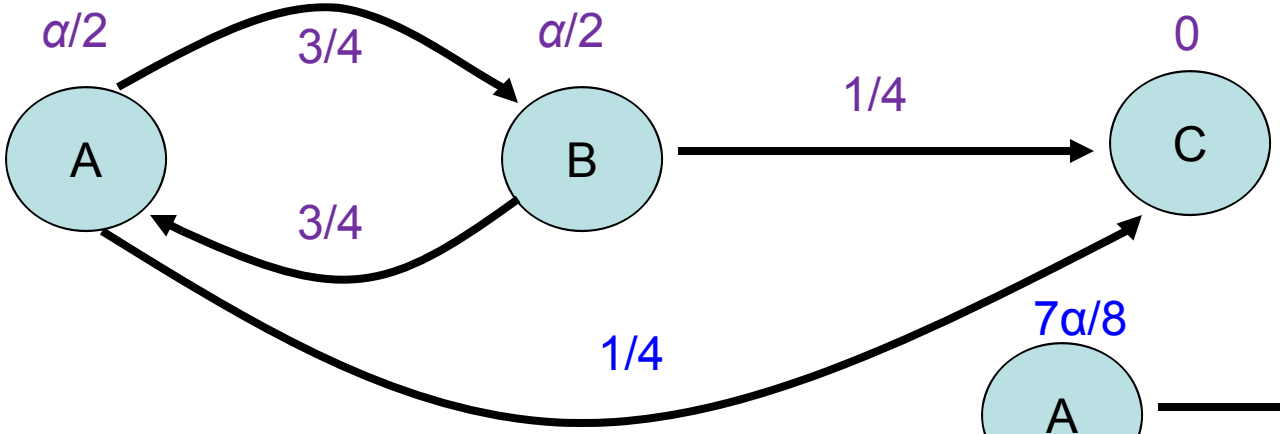
α-exhaustive nature of the Holm's method

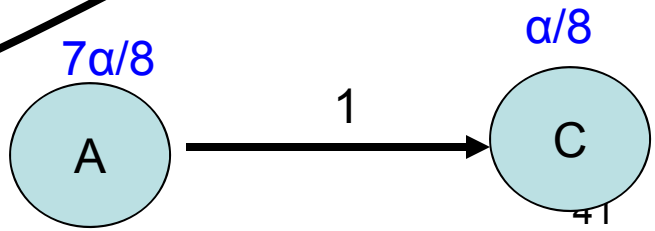




C can be tested only when both *A* and *B* are successful

Truncation of the Holm's test for the primary family



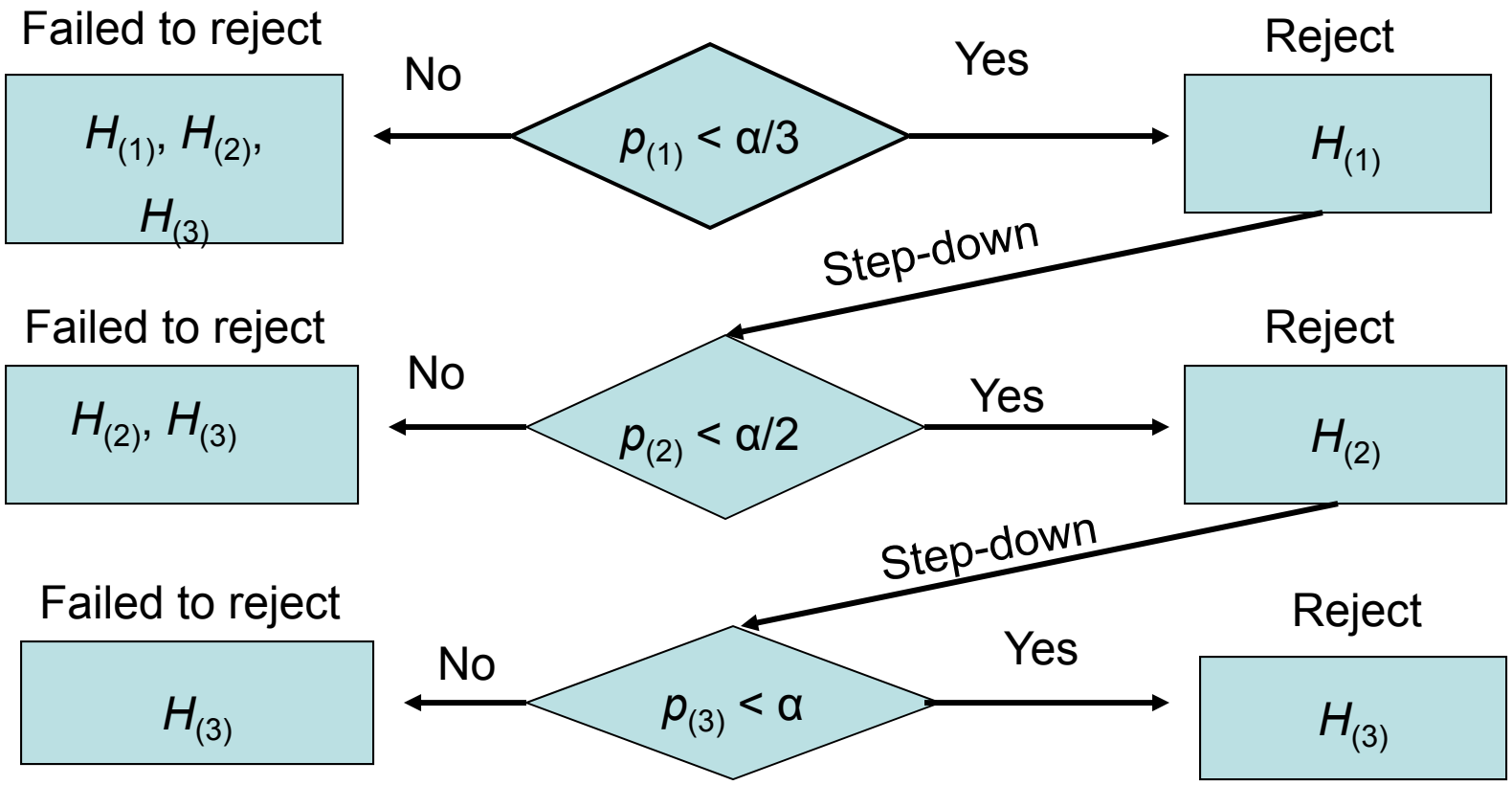


Dmitrienko et al. 2008; Bretz et al. 2009

After B is successful

Regular Holm test with K =3

Ordered p-values of $p_{(1)} \leq p_{(2)} \leq p_{(3)}$
Associated hypotheses $H_{(1)}$, $H_{(2)}$, $H_{(3)}$
Start from the top with the smallest p-value $p(1)$ then step-down



Truncated Holm test for testing *k* hypotheses

(Dmitrienko et al., 2008)

- The truncated Holm test allows passing of alpha from one family to the other, but the calculation of un-used alpha is different than that by the Bonferroni based method
- In the truncated Holm, the critical values for tests are convex combinations of the critical values of the original Holm test and that of the Bonferroni test

$$c_i = \theta(\alpha/(k - i + 1)) + (1- \theta)(\alpha/k), \qquad (i = 1, \ldots, k)$$

where, $0 \le \theta \le 1$ is known as the truncation fraction.

  - At $\theta = 0$, this construct gives the Bonferroni alpha-critical value of α/k.
  - The actual procedure for the truncated Holm remain the same, except that the above new critical values c_i are used

Truncated Holm test for the primary family; $K=2$, $\theta=3/4$

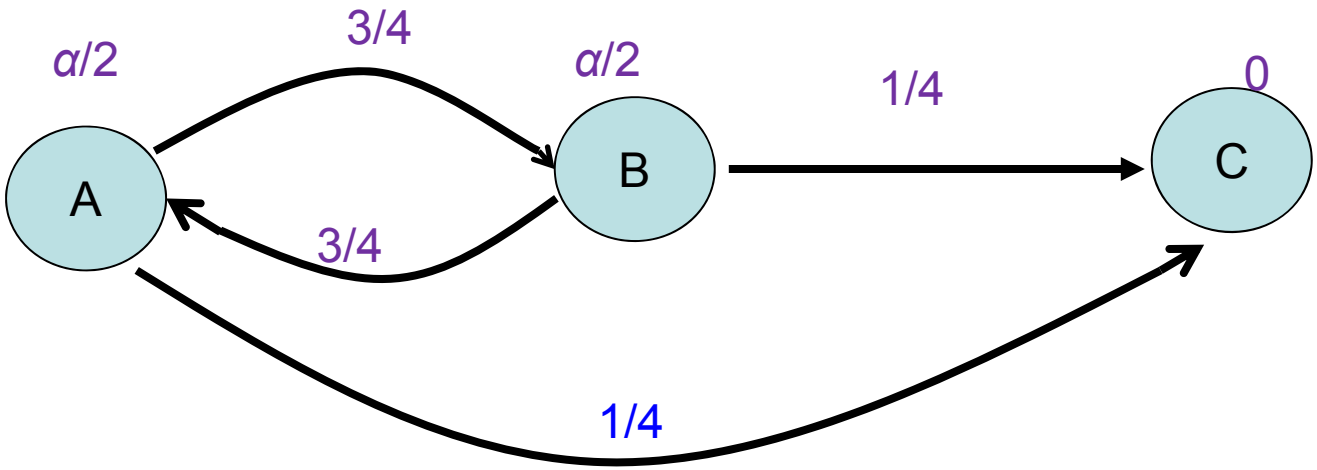
- Family 1 test (assume truncation fraction $\theta=3/4$):

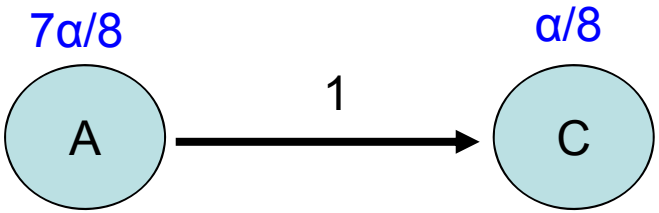
  Reject $H_{(1)}$ if $p_{(1)} < c_1 = \alpha/2$, otherwise, stop testing

  Reject $H_{(2)}$ if $p_{(2)} < c_2 = (1+\theta)\alpha/2 = (7/8)\alpha$ after rejecting $H_{(1)}$, otherwise, stop testing

- Alpha remained for the Family 2 is:
  - All α when in Family 1 all null hypotheses are rejected
  - $\alpha - c_2 = (1 - 7/8)\alpha = (1/8)\alpha$ when in Family 1 $H_{(1)}$ is rejected but $H_{(2)}$ is retained

Graphical representation: truncated Holm test for the primary family; K =2 and θ=3/4





After B is successful

Truncated Holm test for the primary family; K =2 and θ=1/2

- Family 1 test (assume truncation fraction θ =1/2):

  Reject $H_{(1)}$ if $p_{(1)} < c_1 = \alpha/2$, otherwise, stop testing

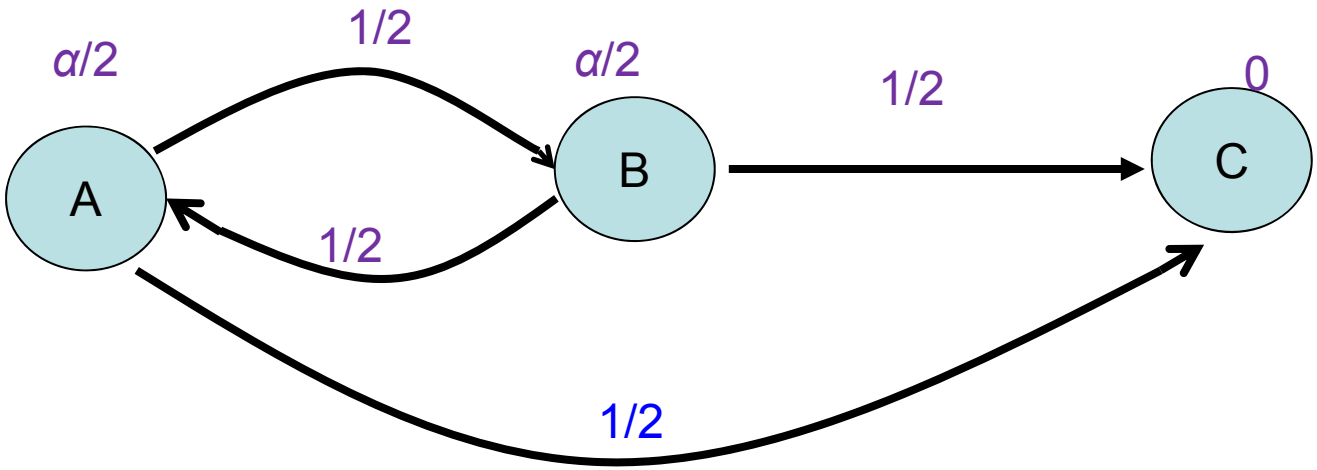
  Reject $H_{(2)}$ if $p_{(2)} < c_2 = (1 + \theta)\alpha/2 = (3/4)\alpha$ after rejecting $H_{(1)}$, otherwise, stop testing

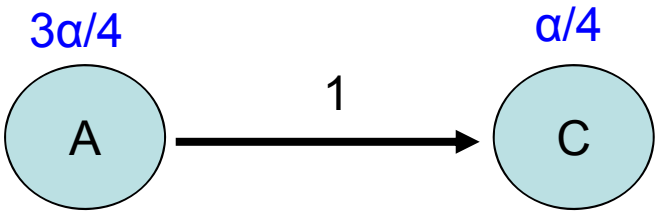
- Alpha remained for Family 2 is:

  All α when in Family 1 all null hypotheses are rejected

  $\alpha - c_2 = (1 - 3/4)\alpha = (1/4)\alpha$ when in Family 1 $H_{(1)}$ is rejected but $H_{(2)}$ is retained

Graphical representation: truncated Holm test for the primary family; K =2 and θ=1/2





After B is successful

Truncated Holm test, K =3 (primary family)

- Family 1 test:
  1) Reject $H_{(1)}$ if $p_{(1)} < c_1 = \alpha/3$, else stop testing
  2) Reject $H_{(2)}$ if $p_{(2)} < c_2 = (\theta + 2)\alpha/6$ after rejecting $H_{(1)}$, else stop testing, and
  3) Reject $H_{(3)}$ if $p_{(3)} < c_3 = (2\theta + 1)\alpha/3$ after rejecting $H_{(1)}$ and $H_{(2)}$.
- Alpha saved for Family 2 is:
  a) All α when in Family 1 all null hypotheses are rejected
  b) $\alpha - 2c_2 = (1 - \theta)\alpha/3$ when in Family 1 $H_{(1)}$ is rejected but $H_{(2)}$ and $H_{(3)}$ are retained
  c) $\alpha - c_3 = 2(1 - \theta)\alpha/3$ when in Family 1 both $H_{(1)}$ and $H_{(2)}$ are rejected but $H_{(3)}$ is retained

An illustrative example

- Consider treatment-to control comparisons on three endpoints in the primary family with the control of alpha at the 0.05 level.
  - Test critical values for the conventional Holm are: 0.05/3, 0.05/2, and 0.05, and those for the equally weighted Bonferroni method are 0.05/3, same for each comparison
  - The endpoint-specific alpha levels for the truncated Holm with a "truncation fraction" of θ =1/2 are*:*

$\alpha_1 = (0.05/3)\,\theta + (0.05/3)(1-\theta) = 0.0167$ (same as 0.05/3)

$\alpha_2 = (0.05/2)\,\theta + (0.05/3)(1-\theta) = 0.0208$ (instead of 0.05/2)

$\alpha_3 = (0.05)\theta + (0.05/3)(1-\theta) = 0.0333$ (instead of 0.05)

An illustrative example (cont'd)

- The unused alphas for passing to secondary family are:

  (i) 0.05 if all three tests are successful

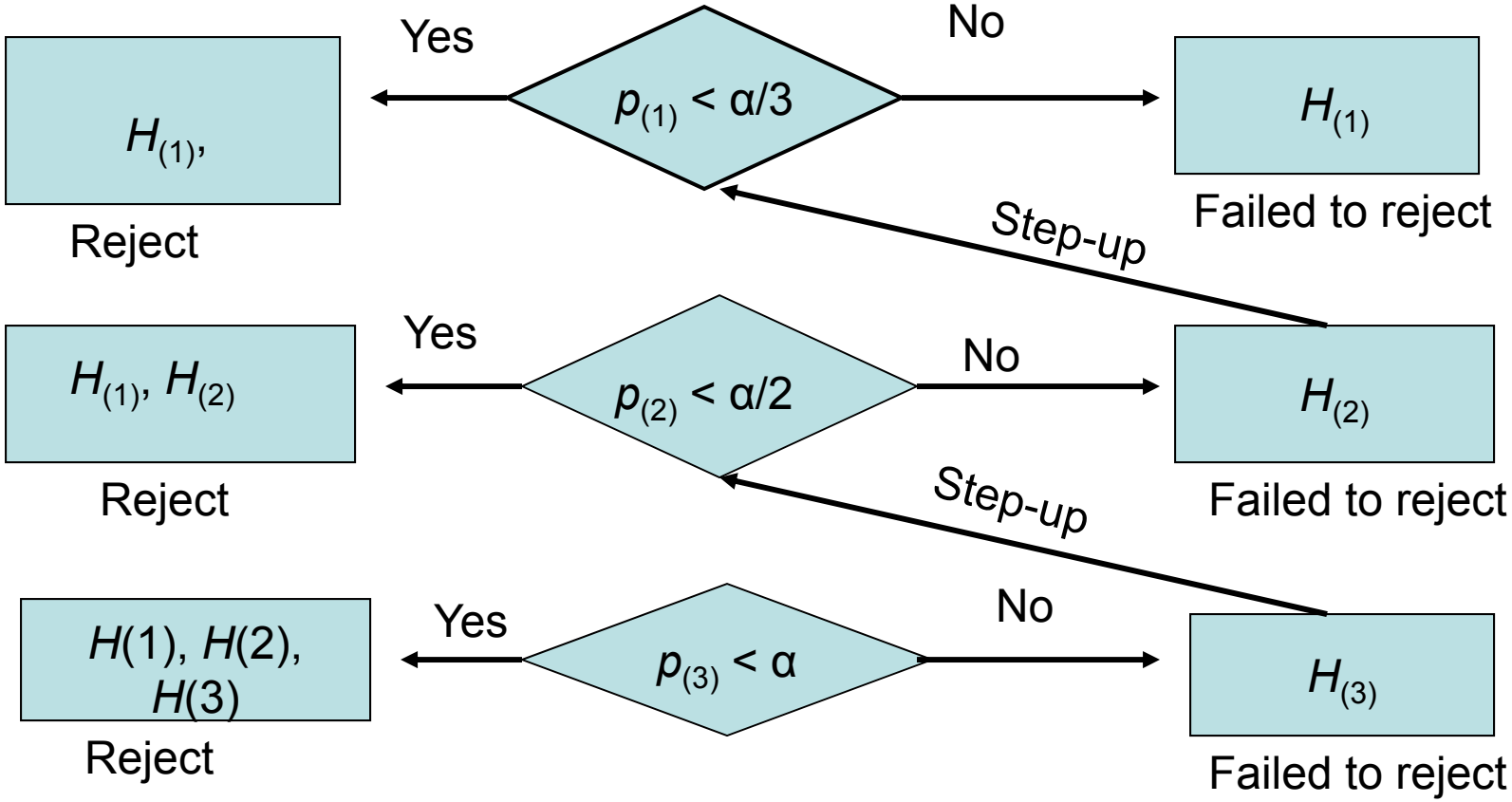
  (ii) $(0.05 - \alpha_3) = 0.05 - 0.0333 = 0.0167$, if the 1st two tests are successful but the last one is not

  (iii) $(0.05 - 2\,\alpha_2) = 0.05 - 2(0.0208) = 0.0084$, if the 1st test is successful but the other two are not.

Hochberg procedure with K =3

Start from the bottom with the largest p-value $p_{(3)}$ then step-up

Ordered p-values of $p_{(1)} \leq p_{(2)} \leq p_{(3)}$
Associated hypotheses $H_{(1)}$, $H_{(2)}$, $H_{(3)}$



Why consider the HP for confirmatory trials?

- Consider for example 1-sided treatment effect p-values of 0.013 and 0.022 on two primary endpoints of a trial designed to compare a new treatment to control.
- One would normally consider such results as acceptable as evidence of treatment effects on the two endpoints, if the procedure employed controls FWER at level 0.025 in the strong sense.
- Thus, if such results can win the trial, then the use of the Bonferroni and the Holm procedures would be unwise, as these procedures would consider such results as statistically not significant and would fail the trial.

Comments for the HP

- **It is NOT assumption-free like the Bonferroni and Holm tests.**
  - It provides adequate FWER control for independent and for certain types of positively correlated tests (Sarkar and Chang, 1997; Sarkar, 1998), but its properties for other types of dependent endpoints are not fully known for more than 2 hypotheses tests .
  - It provides adequate FWER control for testing 2 null hypotheses, when test statistics follow bivariate normal, or bivariate t, or 1-df chi-square distributions with positive correlations.

    References: Sarkar & Chang (1997); Samuel-Cahn (1996); Huque (SIM 2015)

Comments for the HP (cont'd)

- Similar to Holm, HP is α-exhaustive
  - This means that in testing the primary family of null hypotheses of a trial, it is not able to release any alpha for tests for the secondary family of null hypothesis of a trial, unless all null hypotheses in this first family are first rejected.
- However, the truncated HP can be used for the primary family if the desire is also to test the secondary family
- The method of truncation for the HP is the same as that for the Holm (Dmitrienko et al 2008)

Closed testing procedure (CTP)

- Given h elementary hypotheses $H_1, \ldots, H_h$, the CTP considers the 2^h -1 intersection hypotheses:

  $H_{\boldsymbol{F}} = \cap_{i \in \boldsymbol{F}} H_i$ where $\boldsymbol{F}$ stands for $\boldsymbol{I} = \{1, \ldots, h\}$ and all subsets $\boldsymbol{F}$ of $\boldsymbol{I}$

  For example, given h =3 hypotheses H_1, H_2, and H_3,

  $H_{\boldsymbol{F}} \varepsilon \{H_1 \cap H_2 \cap H_3, H_1 \cap H_2, H_1 \cap H_3, H_2 \cap H_3, H_1, H_2, H_3\}$

- Test Procedure:
  - Test each $H_{\boldsymbol{F}}$ at level α or less
  - Reject $H_{\boldsymbol{F}}$ if and only if $H_{\boldsymbol{F}}$ and all higher order interaction hypotheses that include $H_{\boldsymbol{F}}$ are rejected at level α or less

Closed testing procedure (CTP)

- For example, consider 3 hypotheses H_1, H_2, and H_3. Then in order to reject H_1 , one has to reject all the 4 hypotheses:

  $H_1 \cap H_2 \cap H_3$, $H_1 \cap H_2$, $H_1 \cap H_3$ and H_1, testing each at the same significance level α

- Thus, an individual null hypothesis H_i is rejected at level α if and only if every intersection hypothesis H_F that includes H_i (including H_i itself) is rejected at level α.

- Note: for h =3, if for example, H_1 is rejected, then to reject H_2, one has to test only $H_2 \cap H_3$ and H_2, each at level α. Further, if both H_1 and H_2 are rejected, then one has to simply test H_3 at level α

- The CTP strongly controls FWER≤ α (Marcus et al., 1976).

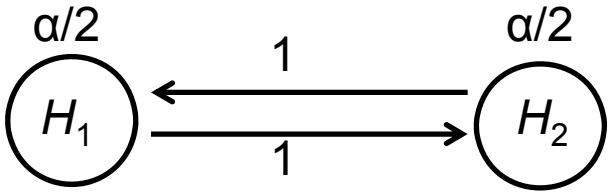
"Consonance" property

- Consonance Property (Gabriel, 1969):
  - The rejection of an intersection hypotheses implies the rejection of at least one of its elementary hypotheses
  - For example, if $H_{\boldsymbol{F}} = \bigcap_{i \in \boldsymbol{F}} H_i$ is rejected at level α, then elementary hypotheses H_i are rejected at level α for at least for one $i \in \boldsymbol{F}$
- Shortcuts to the CTP occur if this property holds (Hommel et al., 2007)

✓ **Sequentially rejective (SR) graphical procedures are (implicitly) related to CTPs that satisfy this property**

CTP with 2 hypotheses and it connection to α-recycling and the graphical method

- Closed testing considers $\{H_1 \cap H_2, H_1 \text{ and } H_2\}$
- Suppose we use Bonferroni test for $H_1 \cap H_2$. That is, reject $H_1 \cap H_2$ if unadjusted $p_j < \alpha/2$ for some $j \in \{1, 2\}$.
- Suppose that $H_1 \cap H_2$ is rejected for $j = 2$. Then by the consonance property of the test for $H_1 \cap H_2$, the hypothesis H_2 is rejected
- Consequently, by the CTP, the test for H_1 is at level α and not at level α/2.
- The above procedure can, therefore, be represented graphically as:



CTP with weighted Bonferroni tests for intersection hypotheses

- Consider h elementary hypotheses $H_1, \ldots, H_h$
- $\boldsymbol{I} = \{1, \ldots, h\}$. For $\boldsymbol{F} = \boldsymbol{I}$ and for any subset $\boldsymbol{F}$ of $\boldsymbol{I}$ consider intersection hypotheses:

  $H_{\boldsymbol{F}} = \cap_{i \epsilon \boldsymbol{F}} H_i$ with weights $w_i(\boldsymbol{F})$ associated with H_i for $i \epsilon \boldsymbol{F}$ so that $\sum_{i \epsilon \boldsymbol{F}} w_i(\boldsymbol{F}) \leq 1$

- Reject $H_{\boldsymbol{F}}$ if $p_i < w_i(\boldsymbol{F})\alpha$ for some $i \varepsilon \boldsymbol{F}$. (Weighted-Bonferroni test for $H_{\boldsymbol{F}}$)
- Example: (next slide)

CTP with weighted Bonferroni tests for intersection hypotheses

- Example: Consider 3 hypotheses H_1, H_2, and H_3

H_F	H_1 w_1	H_2 w_2	H_3 w_3	Reject H_F If
H_{123}	0.6	0.3	0.1	$p_i < w_i(H_{123})\alpha$ for some $i \in \{1, 2, 3\}$
H_{12}	0.8	0.2	-	$p_i < w_i(H_{12})\alpha$ for some $i \in \{1, 2\}$
H_{13}	0.7	-	0.3	$p_i < w_i(H_{13})\alpha$ for some $i \in \{1, 3\}$
H_1	1.0	-	-	$p_1 < \alpha$
H_{23}	-	0.3	0.7	$p_i < w_i(H_{23})\alpha$ for some $i \in \{2, 3\}$
H_2	1	-	-	$p_2 < \alpha$
H_3	-	-	1	$p_3 < \alpha$

Note: H123 = H1 ∩ H2 ∩ H3; H12 = H1 ∩ H2; H13 = H1∩ H3; H23 = H2 ∩ H3
Such a table for CTP was introduced by Dmitrienko et al. (2003)

CTP with Bonferroni tests with weights (BWS) that satisfy consonance

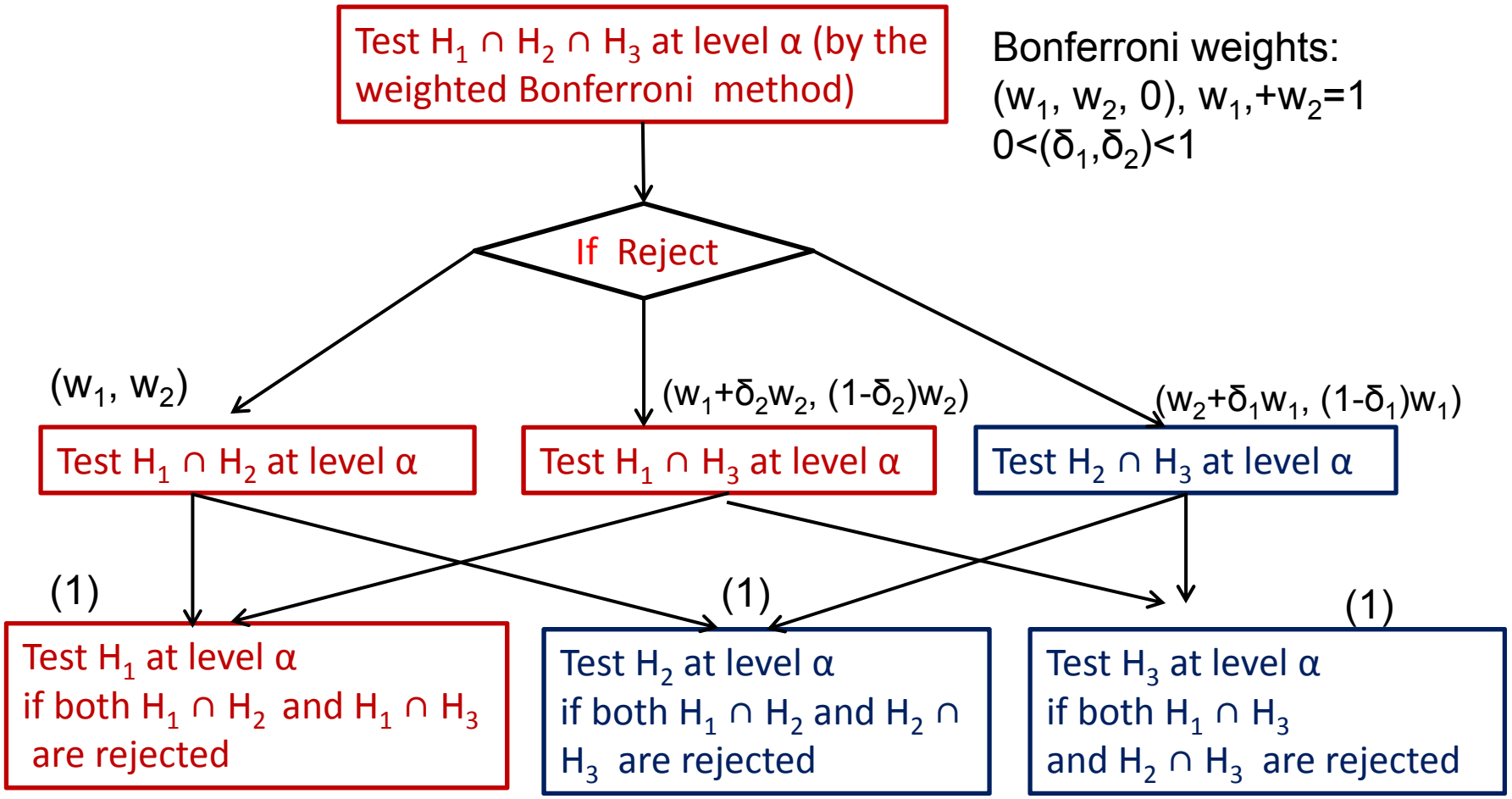
- If in addition, for any intersection hypothesis $H_{F^*} = \cap_{i \in F^*} H_i$, weights $w_i(F^*)$ with $\sum_{i \in F^*} w_i(F^*) \leq 1$ satisfy the following condition

$$w_i(F^*) \geq w_i(F) \text{ for every subset } F^* \text{ of } F \qquad \textbf{(A)}$$

- Note that in the previous weighting scheme this condition is not satisfied

✓ **The Bretz et al. (2009) graphical approach satisfies this condition for all intersection hypotheses H_F .**

Example 3 : H_1 and H_2 are primary hypotheses and H_3 is the secondary

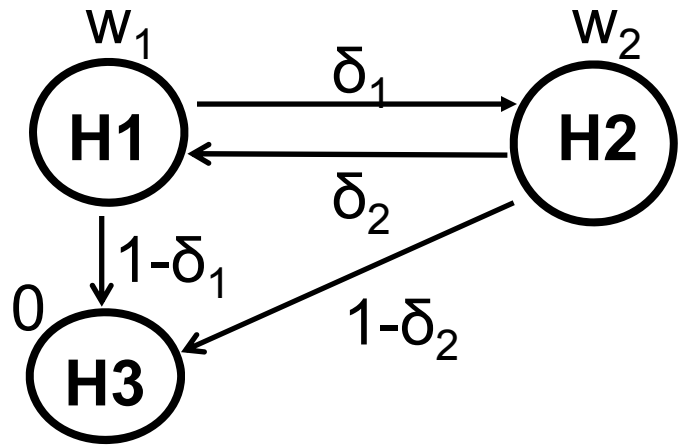


NOTE: H_3 is tested only when at least one primary hypothesis is rejected

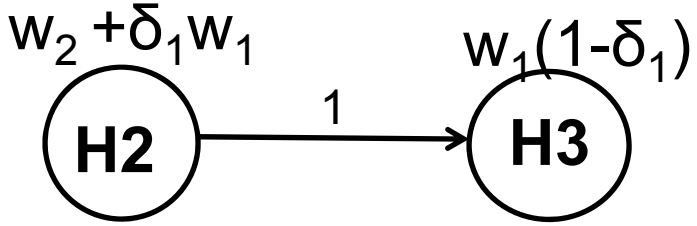
CTP Table (Example 3) with BWs satisfying consonance

Hypotheses	H1	H2	H3
H123	**w_1**	**w_2**	0
H12	w_1	w_2	0
H13	**$w_1+\delta_2 w_2$**	-	**$(1-\delta_2)w_2$**
H1	1	-	-
H23	-	**$w_2+\delta_1 w_1$**	**$(1-\delta_1)w_1$**
H2	-	1	-
H3	-	-	1

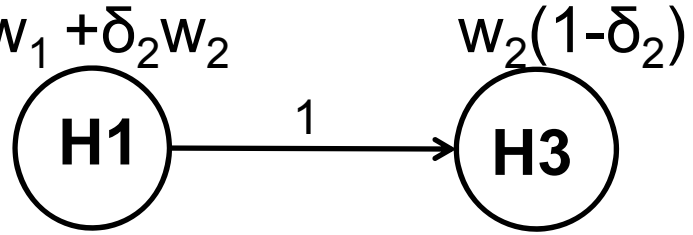
Graphical representation of the CTP Table



Original graph (A)



Graph after rejecting H1 in (A)



Graph after rejecting H2 in (A)

Transition matrix of g-values in (A)

$g_{12} = \delta_1$, $g_{13} = 1-\delta_1$;

$g_{21} = \delta_2$, $g_{23} = 1-\delta_2$; $g_{31} = g_{32} = 0$

Calculation of g-values after rejecting a hypothesis

Transition matrix g-values (original graph):

$g_{12} = \delta_1$, $g_{13} = 1-\delta_1$;

$g_{21} = \delta_2$, $g_{23} = 1-\delta_2$; $g_{31} = g_{32} = 0$

Method for calculating new g-values after rejection of the hypothesis H_j:

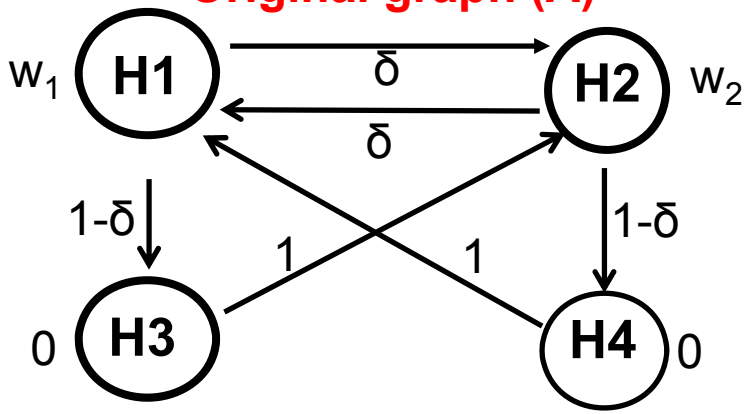
$$g_{\ell k}(\text{new}) = (g_{\ell k}(\text{old}) + g_{\ell j} * g_{jk})/(1-g_{\ell j} * g_{j\ell})$$

Example after rejecting H_1 :

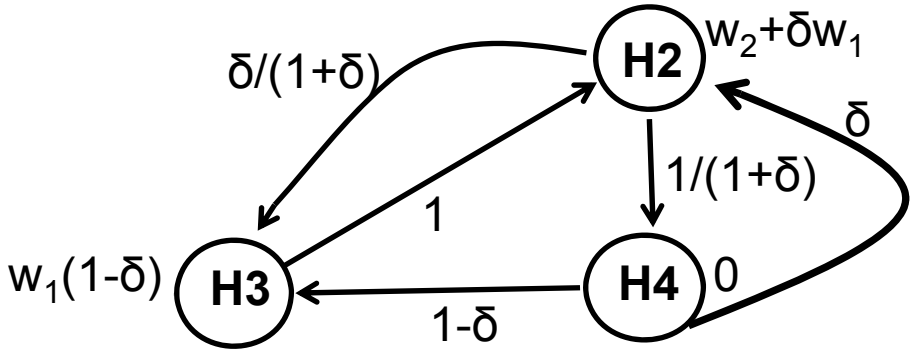
$g_{23} = \{g_{23}(\text{old}) = (1-\delta_2) + g_{23}$ (going through $H_1) = \delta_2(1-\delta_1)\}/\{(1- g_{21} \text{x } g_{21}) = (1-\delta_1\delta_2)\} = 1$

Example 4: (H1,H2) primary, (H3,H4) secondary

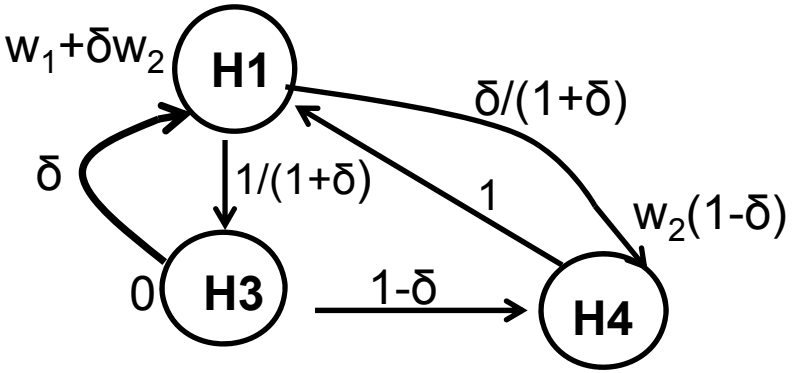
Original graph (A)



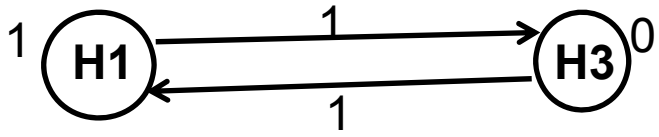
After rejecting H1



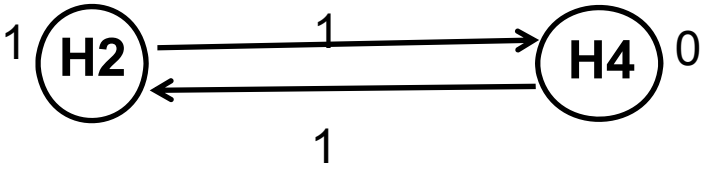
After rejecting H2



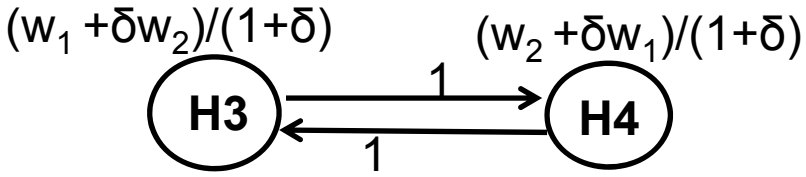
After rejecting H2 and H4



After rejecting H1 and H3



After rejecting H1 and H2



Note: (H1,H3) and (H2,H4) are descending hypotheses pairs (Method: Weighted Bonferroni)

CTP Table with the consonance property (Example 4, using weighted Bonferroni)

Hypotheses	H1	H2	H3	H4
H1234	**w_1**	**w_2**	0	0
H123	w_1	w_2	0	-
H124	w_1	w_2	-	0
H134	**$w_1+\delta w_2$**	-	0	**$w_2(1-\delta)$**
H12	w_1	w_2	-	-
H13	**1**	-	0	-
H14	$w_1+\delta w_2$	-	-	$w_2(1-\delta)$
H1	1	-	-	-
H234	-	**$w_2+\delta w_1$**	**$w_1(1-\delta)$**	0
H23	-	$w_2+\delta w_1$	$w_1(1-\delta)$	-
H24	-	**1**	-	0
H2	-	1	-	-
H34	-	-	**$(w_1 + \delta w_2)/(1+\delta)$**	**$(w_2 + \delta w_1)/(1+\delta)$**
H3	-	-	1	-
H4	-	-	-	1

Weighted parametric tests for greater power

- Weighted parametric methods can be used to increase power of the test procedure whenever the joint multivariate distribution of the test statistics is known. For this case, one can reject the intersection hypothesis

$$H_{\boldsymbol{F}} = \cap_{j\in \boldsymbol{F}} H, \text{ if } \quad p_j \le c_{\boldsymbol{F}} w_j(\boldsymbol{F})\alpha \text{ for some } j\in \boldsymbol{F}$$

where $c_{\boldsymbol{F}} \ge 1$ is the largest constant satisfying

$$\Pr(\mathbf{U}_{j\in \boldsymbol{F}}\{P_j \le c_{\boldsymbol{F}} w_j(\boldsymbol{F})\alpha\} \mid H_{\boldsymbol{F}}) \le \alpha$$

- If the joint multivariate distribution of the test statistics is not fully known, still it is possible to derive conservative upper bounds of the rejection probability for improvements over the Bonferroni approach

Example 5

- Consider a 2-arm targeted subgroup trial design which allocates a proportion $K = 0.5$ of the total trial sample to a targeted subgroup.
  - The interest is to show benefit of the study treatment (in comparison to a control) on a primary and a secondary endpoint either for the overall patient population (OPP) or for the targeted subgroup (TSG).
- Consider 4 test statistics (corresponding to 4 null hypotheses) whose joint distribution is multivariate normal $N_4(\mathbf{0}, \boldsymbol{R})$ which is <u>not</u> fully known

Example 5 (cont'd)

Primary endpoint tests

- Z_0 = test statistic for the OPP
- Z_s = test statistics for the TSG
- (Z_0, Z_s) is bivariate normal with $\rho=\sqrt{K}$ = 0.7071 , when the fraction K in TSG = 0.5

Secondary endpoint tests

- U_0 = test statistic for the OPP
- U_s = test statistics for the TSG
- (U_0, U_s) is bivariate normal with $\rho=\sqrt{K}$ = 0.7071 , when the fraction K in TSG = 0.5

The bivariate joint distribution of 2 test statistics within each family is fully known, but $N_4(\mathbf{0}, \boldsymbol{R})$ across all 4 test statistics are not fully known.

Example 5 (cont'd)

- Consider the intersection hypothesis $H_{\boldsymbol{F}} = \cap_{j\epsilon \boldsymbol{F}} H$, , *and say* $\boldsymbol{F}=\boldsymbol{I} = \{1, 2, 3, 4\}$, then

  $$\Pr(\mathbf{U}_{j\epsilon \boldsymbol{F}}\{P_j \le c_{\boldsymbol{F}} w_j(\boldsymbol{F})\alpha\} | H_{\boldsymbol{F}}) \le$$

  $$\Pr(\mathbf{U}_{j\epsilon\{1,2\}}\{P_j \le c_{\boldsymbol{F}} w_j(\boldsymbol{F})\alpha\} + \Pr(\mathbf{U}_{j\epsilon\{3,4\}}\{P_j \le c_{\boldsymbol{F}} w_j(\boldsymbol{F})\alpha\}$$

- Therefore a conservative $c_{\boldsymbol{F}}$ value for this $\boldsymbol{F}$ can be obtained on setting the above upper bound to α. For this example, this $c_{\{1,2,3,4\}} = 1.1754$
- Similarly, for $\boldsymbol{F}= \{1, 2, 3\}$, its $c_{\{1,2,3\}}$ can be conservatively obtained for this example from the equation:

  $$\Pr(\mathbf{U}_{j\epsilon\{1,2\}}\{P_j \le c_{\boldsymbol{F}} w_j(\boldsymbol{F})\alpha\} + \Pr(P_3 \le c_{\boldsymbol{F}} w_j(\boldsymbol{F})\alpha\} = \alpha$$

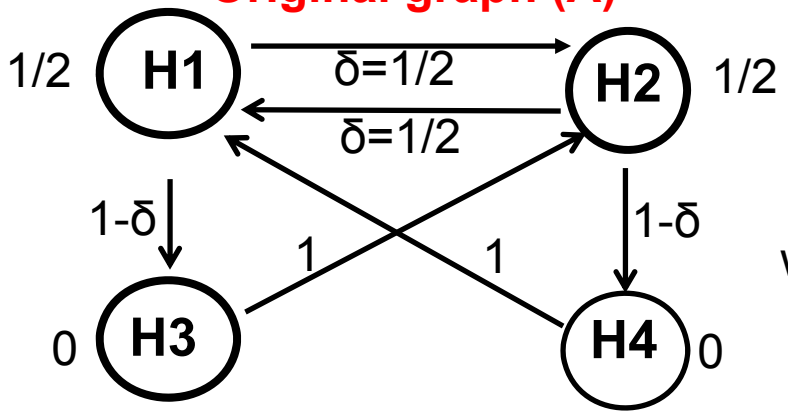
- Therefore, one can construct the CTP table with $w_1=w_2=1/2$, and $w_3=w_4=0$, as in the next slide

CTP Table with the consonance property (Example 5)

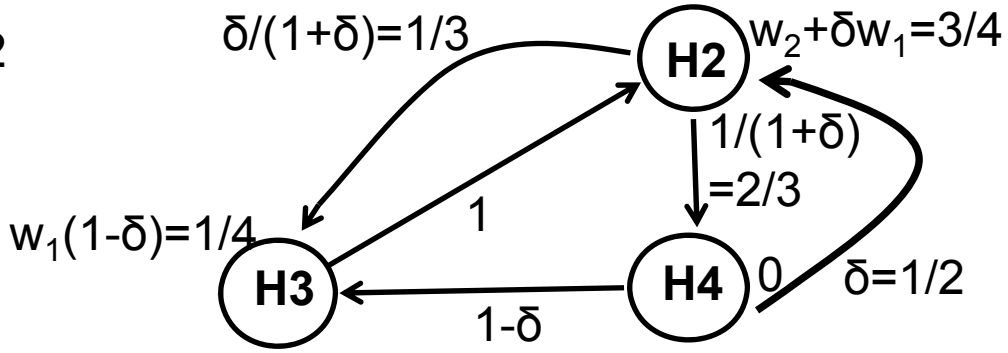
Hypo theses	H1	H2	H3	H4
H1234	0.5877	0.5877	0	0
H123	0.5877	0.5877	0	-
H124	0.5877	0.5877	-	0
H134	0.75	-	0	0.25
H12	0.5877	0.5877	-	-
H13	1	-	0	-
H14	0.75	-	-	025
H1	1	-	-	-
H234	-	0.75	0.25	0
H23	-	0.75	0.25	-
H24	-	1	-	0
H2	-	1	-	-
H34	-	-	0.5877	0.5877
H3	-	-	1	-
H4	-	-	-	1

Example 5: Graphical representation

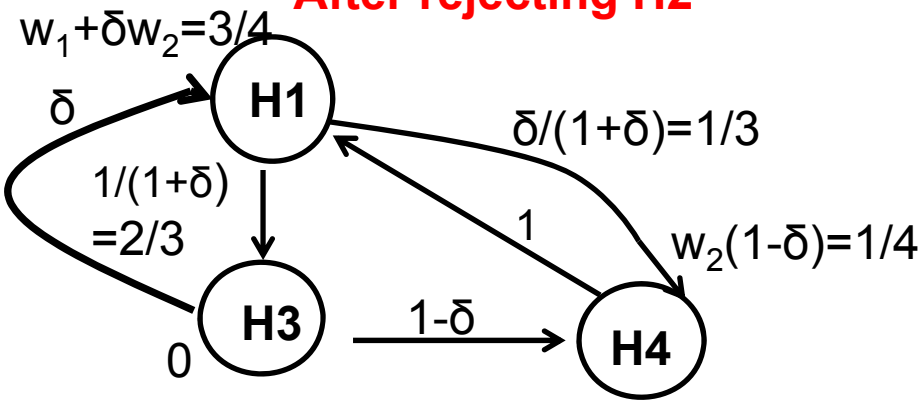
Original graph (A)



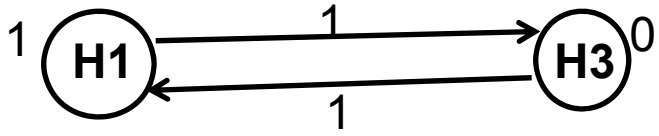
After rejecting H1



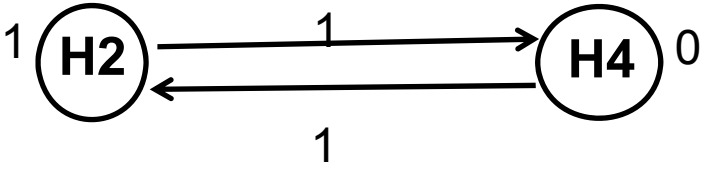
After rejecting H2



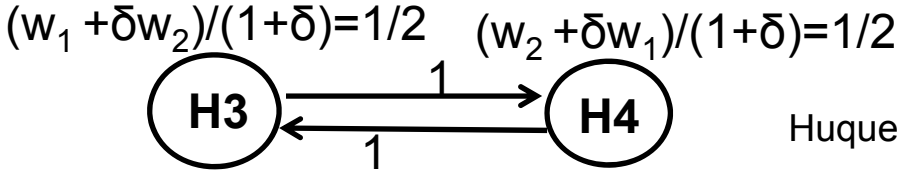
After rejecting H2 in H4



After rejecting H1 and H3



After rejecting H1 and H2



Testing for the parametric approach (Bretz et al., 2011) and cautionary remark

- **Testing** (graphical approach):
  (1) Start with $\boldsymbol{F}=\boldsymbol{I}$ (considering all 4 null hypotheses) and find $c_{\boldsymbol{F}}$
  Reject H_j if $p_j \le c_{\boldsymbol{F}} w_j(\boldsymbol{F})\alpha$. Suppose H_j is rejected.
  (2) Descend to $\boldsymbol{F} = \boldsymbol{I} \setminus \{j\}$ after rejecting H_j
  Reject H_i if $p_i \le c_{\boldsymbol{F}} w_i(\boldsymbol{F})\alpha$, for $i \epsilon \boldsymbol{F} = \boldsymbol{I} \setminus \{j\}$
  (3) Continue descending as in (2) till there is no rejection
- **Caution**: The above sequentially-rejective graphical approach is not valid for all δ. For Example 3, consonance property fails for $0 \le \delta < 0.1754$. However, the CTP is valid for all δ.

Sequentially rejective graphical approach using Simes tests for greater power

- For testing any intersection hypothesis $H_F = \cap_{j \varepsilon F} H$, weighted Simes test is uniformly more powerful than the corresponding weighted Bonferroni test.
  - If H_J is rejected by the weighted Bonferroni test then it is also rejected by the Simes test; the latter rejects more hypotheses
- However, CTP with Simes test does not satisfy the consonance property, as such, the usual sequentially rejective graphical approach is not possible.

Sequentially rejective graphical approach using Simes tests (cont'd)

- Nonetheless, Bretz et al. (2011) show that CTP with Simes test partially satisfies the consonance property.
- Consequently, they propose a 2 step procedure:
  - The first step uses the weighted Bonferroni based sequentially rejective graphical method for rejecting hypotheses that can be rejected by this method.
  - The second step is then uses the sequentially rejective graphical method on the remaining non-rejected hypotheses using the weighted Simes test with the weights originally assigned to these non-rejected hypotheses by the Bonferroni procedure.

Caveats for the Simes test

- Simes test is not assumption free and raises issues for clinical trial applications
- Sarkar (1998), Sarkar & Chang (1997) work show that
  - Simes test is a valid test if the joint distribution of the test statistics follow a standard multivariate normal with all correlations equal and non-negative.
  - It is also a valid test if the joint distribution of the test statistics that follow a multivariate t-distribution of Dunnett and Sobel (1954)
  - It is also a valid test for chi-square tests for the above normal distribution
- Various simulation results seem show that Simes test is also a valid test for multivariate normal with non-negative correlations

Concluding Remarks - Part I

- For confirmatory trials, statistical approaches need to consider their hierarchical structures of test hypotheses and their families for gaining efficiency and optimizing power for the primary hypotheses
- In these approaches, for making conclusions at the individual hypotheses levels, strong sense FWER control is needed across both the primary and secondary families of hypotheses.
- Two key statistical approaches for this have been developed that apply to confirmatory clinical trials

  a) Gatekeeping approaches (see FDA tutorial, 2014 BASS)

  b) SR graphical methods (topic of this session)

Concluding Remarks - Part I (cont'd)

- Both the gatekeeping and graphical approaches can handle the following two cases:
  a) Primary hypotheses tests do not depend in any way on the results of the secondary hypotheses test results
  b) Primary hypotheses can be re-tested on recycling some alpha from the test results of the secondary hypotheses
- Both approaches for greater power can account for correlations between endpoints with some modifications when these correlations can be pre-specified.
- The graphical method is (implicitly) related to the CTP with Bonferroni weights.

Outline (Part II)

- Brief introductions on *B*-values and *Z*-scores used in GS (Group Sequential) test procedures and on α-spending functions (Ref: Proshan, Lan and Wittes; 2007)
- **GS test procedures for testing multiple hypotheses**
  - **Methods based on the Bonferroni inequality**
  - **Method based on the CTP (Tang & Geller, 1999)**
  - **The case of testing 2 hypotheses**
  - **The general case of testing multiple hypotheses on using the graphical method (Maurer & Bretz, 2013)**
- Example of a GS trial design for testing a primary and a secondary endpoint of a trial
- Concluding remarks (Part II)

B-values and Z-scores

- Consider a 2-arm trial which is designed with a total sample size of N subjects per treatment arm
- Let S_{n1} = the sum statistic for treatment difference at an interim look #1 based on a sample size of n_1 subjects per treatment arm
- Define: $B(t_1) = S_{n1}/(V_N)^{1/2}$ where $V_N = \text{Var}(S_N) = 2N\sigma^2$
- Then

$$\text{Var}\{B(t_1)\} = n_1/N = t_1 \text{ (information fraction at look \#1)}$$

$$Z(t_1) = (S_{n1}/\sqrt{V_{n1}}) = (S_{n1}/\sqrt{V_N})\,(V_N/V_{n1})^{1/2} = B(t_1)/(t_1)^{1/2}$$

(because $\text{Var}(S_{n1}) = 2n_1\sigma^2$ and $V_N/V_{n1} = 1/t_1$)

B-values and Z-scores

- Consider now the 2nd look with sample size of

  $n_2 = n_1 + r$ (per treatment arm)

- Then

  $B(t_2) = S_{n2}/\sqrt{V_N} = (S_{n1} + S_r)/\sqrt{V_N}$

- Consequently,

  $\text{Var}\{B(t_2)\} = t_2$, $\text{Cov}\{B(t_1), B(t_2)) = t_1$ and

  $\text{Corr}\{B(t_1), B(t_2)\} = (t_1/t_2)^{1/2}$ for $t_1 \leq t_2$

  $\text{Corr}\{Z(t_1), Z(t_2)\} = (t_1/t_2)^{1/2}$ for $t_1 \leq t_2$

Joint distributions

- Given $t_1 \le t_2 \le \ldots \le t_k$, if assume that $B(t_1)$, $B(t_2)$, …, $B(t_k)$ jointly follow a multivariate normal distribution, then

$$E\{B(t_j)\}=0 \text{ under } H_0;\ \text{Cov}\{B(t_i),B(t_j)\} = t_i \text{ for } t_i \le t_j$$

- Also, $Z(t_j) = B(t_j)/t_j^{1/2}$ is the Z-score corresponding to $B(t_j)$
- Further, $Z(t_1)$, $Z(t_2)$, …, $Z(t_k)$ jointly follow a multivariate normal distribution with

$$E\{Z(t_j)\}=0 \text{ under } H_0;\ \text{Cov}\{Z(t_i), Z(t_j)\}= (t_i/t_j)^{1/2} \text{ for } t_i \le t_j$$

Expected values of $B(t_j)$ and $Z(t_j)$

- $E\{B(t_j)\} = n_j\delta/\{2N\sigma^2\}^{1/2} = (n_j/N)\{(N/2)^{1/2}\delta/\sigma\}$
  $= t_j\theta$, where $\theta = (N/2)^{1/2}\delta/\sigma$
- This θ is usually called the drift parameter

  Note that for a fixed sample trial design $\theta = Z_{1-\alpha} + Z_{1-\beta}$.
  For example, if $\alpha = 0.025$ and power = 90%, then $\theta = 3.2415$

- $E\{Z(t_j)\} = E\{B(t_j)\}/(t_j)^{1/2} = (t_j)^{1/2}\,\theta$
- The book (by Proshan, Lan Wittes; 2007) shows how the different GS methods use *B* and Z-statistics and their distributions to set-up GS-boundaries.

α-spending functions

- Lan and DeMets (1983) introduced the concept of α-spending functions. They showed methods for construction GS boundaries that do not require pre-specifying the number or timing of the looks.
- Any non-decreasing function $f(\alpha, t)$ in the information time t $(0 \le t \le 1)$ parametrized by the overall significance level α can be an α-spending function if it satisfies the following conditions:
  1) $f(\alpha, t) \le f(\alpha, t') \text{ for } 0 \le t < t' \le 1$
  2) $f(\alpha, 0) = 0$
  3) $f(\alpha, 1) = \alpha$

Examples of α-spending functions

- OF-like:

  $f_1(\alpha, t) = 2[1-\Phi(z_{1-\alpha/2}/t^{1/2})]$,

  where $z_{1-\alpha/2}$ is the deviate on the standard normal density curve so that area under the curve to the tight of it is $\alpha/2$

- Linear: $f_2(\alpha, t) = \alpha t$
- PK-like:

  $f_3(\alpha, t) = \alpha \log_e\{1+(e-1)t\}$

- Hwang-Shih-Decani (1990):

  $f_4(\alpha, t) = \alpha\{1- \exp(-\lambda t)\}/\{1- \exp(-\lambda)\}$

Calculation of boundary values using the OF-like α-spending function

- Given an α-spending function, one needs to find the nominal significance level $\alpha_t(\alpha)$ at information time t so that H_0 is rejected when p_t at information time t is smaller than $\alpha_t(\alpha)$. We show as an example how to find this for the OF-like alpha spending function.
- Suppose that $\alpha = 0.025$ and the 1st look occurs at $t_1 = 0.30$.
  - We spend $f_1(\alpha, 0.30) = 2[1-\Phi(z_{1-\alpha/2}/(0.30)^{1/2})]= 2[1-\Phi(2.2414027/(0.30)^{1/2})]= 0.0000427$.
  - Therefore, critical value $C_1= 3.9285725$ form $\Pr(Z(t_1) > C_1) = 0.0000427$.
  - We reject H_0 if $p_1 > \alpha_1(\alpha) = 0.0000427$ or $Z(t_1) > C_1= 3.9285725$

Calculation of boundary values using an α-spending function (cont'd)

- Suppose the H_0 is not rejected at the 1st look and the 2nd look occurs at $t = 0.65$.
- The cumulative type I error rate by $t = 0.65$ is
  - $f_1(\alpha, 0.65) = 2[1-\Phi(z_{1-\alpha/2}/(0.65)^{1/2})]= 2[1-\Phi(2.2414027/(0.65)^{1/2})]= 0.0054339$.
- We determine the boundary C_2 by solving the equation: $\Pr\{(Z(t_1) > 3.9285725) \cup (Z(t_2) > C_2)\} = 0.0054339$.
  - Therefore, $C_2 = 2.5479$ and $\alpha_2(\alpha) = 0.0054187$
- We reject H_0 if $p_2 > \alpha_2(\alpha) = 0.0054187$ or $Z(t_2) > C_2 = 2.5479$

Calculation of boundary values using an α-spending function (cont'd)

- Suppose that H_0 is not rejected at the 2nd look and the trial moves to the final look at $t = 1$
- The cumulative type I error rate by $t = 1$ is 0.025
  - $f_1(\alpha, 1) = 2[1-\Phi(z_{1-\alpha/2}]= \alpha = 0.025$.
- We determine the boundary C_3 by solving the equation: $\Pr\{(Z(t_1) > 3.9285725)\cup(Z(t_2) > 2.5479)\cup(Z(t_3) > C_3)\} = 0.025$.
- Therefore, $C_3 = 1.9897$ and $\alpha_3(\alpha) = 0.023312$
- We reject H_0 if $p_3 > \alpha_3(\alpha) =0.0233$, or $Z(t_3) > C_3= 1.9897$.

A general recursive formula for calculations of c_j and α_j values

$$f(\alpha, t_j) = \Pr\left\{\bigcup_{i=1}^{j} Z(t_i) > c_i\right\}$$

$$= f(\alpha, t_{j-1}) - \Pr\left\{\bigcap_{i=1}^{j-1} Z(t_i) \le c_i \bigcap Z(t_j) > c_j\right\}$$

Free software for calculations from: ww.medsch.wisc.edu/landemets/
There are other software, e.g., East 6.3

Results from the East 6.3 software on using the OF-like α-spending function

Overall α = 0.025

Look #	Information fraction	Cumulative α spent	Efficacy boundary
1	0.30	0.00004	0.00004
2	0.65	0.00543	0.00542
3	1.00	0.025	0.02331

Overall α = 0.0125

Look #	Information fraction	Cumulative α spent	Efficacy boundary
1	0.30	0.00001	0.00001
2	0.65	0.00194	0.00194
3	1.00	0.0125	0.01188

GS testing for multiple endpoints of clinical trials

- Methods based on the Bonferroni inequality
- CTP based procedure (Tang & Geller, 1999)
- Alpha-recycling method for the case of testing 2 hypotheses
- Alpha-recycling method for the general case using the graphical approach (Maurer & Bretz, 2013)

Methods based on the Bonferroni inequality

1) Given h hypotheses $H_1, \ldots, H_h$, assign significance level of α_j for each H_j so that the sum

$$\alpha_1 + \alpha_2 + \ldots + \alpha_h = \alpha.$$

Then apply univariate GS testing method to each H_j (Jennison & Turnbull, 2000)

2) Pre-specify Bonferroni based rejection boundary α_t for $t = 1, \ldots k$, so that the sum

$$\alpha_1 + \alpha_2 + \ldots + \alpha_k = \alpha$$

Then at each time t apply a multiple testing procedure to h hypotheses. The resulting procedure protects FWER over all hypotheses and time points. (Maurer & Bretz, 2013)

CTP based test procedure (Tang & Geller, 1999)

- Consider testing h hypotheses, and let $\boldsymbol{I} = \{1, \ldots, h\}$. Consider

  $\boldsymbol{F} = \boldsymbol{I}$ or any non-empty subset of $\boldsymbol{I}$, and

  $H_{\boldsymbol{F}} = \cap_{j \in \boldsymbol{F}} H_j$ be the intersection hypothesis, i.e., treatment difference $\delta_j = 0$, for $j \epsilon \boldsymbol{F}$

- Consider a group sequential trial with k looks at information times $t = t_1, \ldots, t_k$

CTP based test procedure (Tang & Geller, 1999)

- Let

  Z_F = test statistic used for testing H_F

  $Z_{F,t}$ = test statistic Z_F calculated at information time t

  $c_{F,t}$ for ($t = t_1, \ldots, t_k$) are one-sided GS boundary values for testing H_F, determined so that

  $$\Pr\{Z_{F,t} > c_{F,t} \text{ for some } t \mid H_F\} \leq \alpha$$

- The GS test procedure can then be stated as in the next slide

CTP based test procedure (Tang & Geller, 1999)

- ***Step* 1:** Conduct interim analyses to test H_I based on the group sequential boundary $\{c_{I,t}, t = t_1, \ldots, t_k\}$ using $Z_{I,t}$
- ***Step* 2:** When H_I is rejected, say at time $t = t^*$, apply the CTP to test the other hypotheses H_F using Z_{F,t^*} with c_{F,t^*} as the critical value
- ***Step* 3:** If any hypothesis is not rejected, continue the trial to the next stage, in which the closed testing is repeated (with the previously rejected hypotheses automatically considered rejected w/o retesting)
- ***Step* 4**: Reiterate ***Step* 3** until all hypothesis are rejected or the trial reaches the last stage k

Tang & Geller procedure simplifies when using BWs that satisfy consonance property

- If the CTP with Bonferroni weights satisfies consonance property and the alpha-spending function satisfies certain condition, then the Tang & Geller CT based procedure enjoys certain key benefits:

  a) Allows construction of SR (graphical) testing procedures which lead to recycling of alpha form one hypothesis to another in a manner as shown in Maurer and Bretz (2013)

  b) Existing software can be used to finding nominal significance level for the test of each hypothesis at each interim look so that FWER of the procedure is controlled at level α (e.g., for α = 0.025 for 1-sided tests)

Recall: Test for intersection hypotheses with Bonferroni weights in the CTP

- Consider intersection hypotheses $H_{\boldsymbol{F}} = \cap_{i \in \boldsymbol{F}} H_i$,

  where $\boldsymbol{F} = \boldsymbol{I} =$ {1, …, h} or $\boldsymbol{F}$ is a subset of $\boldsymbol{I}$

  and h is number of hypotheses tested

  Assign weights $w_i(\boldsymbol{F})$ for $i \epsilon \boldsymbol{F}$ so that $\sum_{i \epsilon \boldsymbol{F}} w_i(\boldsymbol{F}) \leq 1$

  Reject $H_{\boldsymbol{F}}$ if $p_i < w_i(\boldsymbol{F})\alpha$ for some $i \epsilon \boldsymbol{F}$.

- Consonance property is satisfied if in addition, for any intersection hypothesis $H_{\boldsymbol{F}^*} = \cap_{i \epsilon \boldsymbol{F}^*} H_i$, with weights $w_i(\boldsymbol{F}^*)$ and $\sum_{i \epsilon \boldsymbol{F}^*} w_i(\boldsymbol{F}^*) \leq 1$ satisfy the following condition

  $w_i(\boldsymbol{F}^*) \geq w_i(\boldsymbol{F})$ for every subset $\boldsymbol{F}^*$ of $\boldsymbol{F}$

First consider the case of testing 2 hypotheses in a non-GS setting

- CTP considers hypotheses: $H_1 \cap H_2$, H_1 and H_2, one intersection and two singleton hypotheses
- Consider $H_{\boldsymbol{F}} = H_1 \cap H_2$, $\boldsymbol{F} = \{1,2\}$. Assign weights:

  $w_1(\boldsymbol{F}) = 0.8$, and $w_2(\boldsymbol{F}) = 0.2$, so that

  $w_1(\boldsymbol{F})\alpha = 0.8 \times 0.025 = 0.02$, and $w_2(\boldsymbol{F})\alpha = 0.2 \times 0.025 = 0.005$
- Consonance property is satisfied; because, if $H_{\boldsymbol{F}}$ is rejected then each of the two singleton hypothesis will be tested with weight =1, i.e., at the full significance level of 0.025.
- SR procedure applies, for example, if $p_i < w_i(\boldsymbol{F})\alpha$ for $i = 1$, then H_1 is rejected; consequently, and H_2 can be tested at the full significance level, i.e., there is recycling of alpha from H_1 to H_2.

Case of testing 2 hypotheses in the GS setting

- Suppose that the first interim look is at $t = t1$ at which time the unadjusted p-values are $p_{1,t1}$ and $p_{2,t1}$, then one would reject $H_{\boldsymbol{F}} = H_1 \cap H_2$, where $\boldsymbol{F} = \{1,2\}$, if

$$p_{1,t1} < \alpha_{1,t1}\,(w_1(\boldsymbol{F})\alpha,\ t=t1) \quad \text{or } p_{2,t1} < \alpha_{2,t1}\,(w_2(\boldsymbol{F})\alpha,\ t=t1)$$

- Where the boundary critical value $\alpha_{1,t1}$ is now obtained using the alpha-spending function $f_1(\gamma, t)$ at $f_1(\gamma = w_1(\boldsymbol{F})\alpha, t=t1)$. Similarly, $\alpha_{2,t1}$ is obtained using the spending function $f_2(\gamma, t)$ at $f_2(\gamma = w_2(\boldsymbol{F})\alpha, t=t1)$.

Case of testing 2 hypotheses in the GS setting (cont'd)

For example, if $w_1(\boldsymbol{F}) = 0.8$, $w_2(\boldsymbol{F}) = 0.2$, and $\alpha = 0.025$. Then from the OF-like spending function at $t1=0.30$,

$$\alpha_{1,t1} = 0.00002 \quad \text{and} \quad \alpha_{2,t1} = 2.977\text{E-}07$$

- Suppose that $H_{\boldsymbol{F}}$ is not rejected at t =t1 then one proceeds to the interim look at $t=t2$ (e.g., $t2 = 0.65$).

Case of testing 2 hypotheses but in the GS setting (cont'd)

- At the second interim look t =t2 one would similarly calculate unadjusted p-values are $p_{1,t2}$ and $p_{2,t2}$, and would reject $H_{\boldsymbol{F}} = H_1 \cap H_2$, where $\boldsymbol{F}$ ={1,2} if

$$p_{1,t2} < \alpha_{1,t2}(w_1(\boldsymbol{F})\alpha, t=t2) \quad \text{or } p_{2,t2} < \alpha_{2,t2}(w_2(\boldsymbol{F})\alpha, t=t2)$$

- Where the boundary critical value $\alpha_{1,t2}$ is now obtained using the spending function $f_1(\gamma, t)$ at $f_1(\gamma = w_1(\boldsymbol{F})\alpha, t=t2)$. Similarly, $\alpha_{2,t2}$ is obtained using the spending function $f_2(\gamma, t)$ at $f_2(\gamma = w_2(\boldsymbol{F})\alpha, t=t2)$.

Case of testing 2 hypotheses but in the GS setting (cont'd)

- For example, with the same $w_1(\boldsymbol{F}) = 0.8$, $w_2(\boldsymbol{F}) = 0.2$, and $\alpha = 0.025$, from the O-F-like spending function at $t2$=0.65,

$$\alpha_{1,t2} = 0.0039 \quad \text{and } \alpha_{2,t2} = 0.000498$$

- Suppose that $H_{\boldsymbol{F}}$ is now rejected, say, at t =t2 for i = 2, where $i \varepsilon \boldsymbol{F}$ ={1,2}. Then things happens

Case of testing 2 hypotheses but in the GS setting (cont'd)

- Since H_F is rejected at t =t2 for i = 2, then because of consonance of the CTP H_2 is rejected. Therefore, at t =t2, CTP allows testing H_1 at the updated level $\alpha^*_{1,t2}(\alpha)$ with the transfer of weight of $w_2(F)$ at H_2 to H_1 with the total weight at H1 being $w_2(F) + w_1(F) = 1$.

- Consequently, $\alpha^*_{1,t2}(\alpha) = 0.00542$ is now obtained using the α-spending function $f_1(\gamma, t)$ at $f_1(\gamma = \alpha, t = t2)$.

- Thus, there is recycling of alpha similar to that for the non-GS setting, but for the GS setting, it occurs through the α-spending function from one hypothesis to the another if one of them is rejected

Case of testing 2 hypotheses but in the GS setting (cont'd)

- Suppose now that H_1 at t =t2 when tested at level $\alpha^*_{1,t2}$ (α) is not rejected, then one would proceed to interim look t = t3 to test H_1 with the assumption that H_2 remains rejected at t3.

- Therefore, at t3 (i.e., the final look), H_1 would be tested at level $\alpha^*_{1,t3}$ (α) = 0.02331

Key points

- Note that in the previous slide at t =t2 after H_2 is rejected, H_1 is tested at level $\alpha^*_{1,t2}$ (α), and not at level α.

  After success on one hypothesis, wrongfully testing the other hypothesis at the full level α can inflate the FWER.

  Instead, one needs to calculate and use $\alpha^*_{1,t2}$ (α) by a standard software such as East 6.3 using a pre-specified α-spending function

- The spending function applied needs to satisfy the following condition:

  The difference function $f(\gamma, t_j) - f(\gamma, t_{j-1})$ is monotonically non-decreasing in γ for $j = 1, \ldots, k$

  OF-like alpha spending function satisfies this condition

Graphical algorithm by Maurer & Bretz (2013) for the general case

- Set $t = t_1$ (1st interim look), $\boldsymbol{I} = \{1, 2, \ldots, h\}$, and weights $w_i(\boldsymbol{I})$ for $i \epsilon \boldsymbol{I}$

1) At interim look t, calculate p-values $p_{i,t}$ and boundary critical values $\alpha_{i,t}$ for $i \epsilon \boldsymbol{I}$ on using alpha of $w_i(\boldsymbol{I})\alpha$

2) Find a $j \epsilon \boldsymbol{I}$ such that H_j is rejected on observing $p_{j,t} < \alpha_{j,t}$; go to Step-3. If no such j exists and $t < t_k$, then go to Step-1 but at $t = t_u$ (the next look, $u = 1, .., k$)

3) Update the graph:

Graphical algorithm by Maurer & Bretz (2013) for the general case (cont'd)

3) Update the graph:

$\boldsymbol{I} = \boldsymbol{I} \setminus \{j\}$

New $w_\ell(\boldsymbol{I}) = w_\ell(\boldsymbol{I}) + w_j(\boldsymbol{I}) * g_{j\ell}$ for $\ell \in \boldsymbol{I}$; zero otherwise

$g_{\ell k}(\text{new}) = (g_{\ell k}(\text{old}) + g_{\ell j} * g_{jk})/(1 - g_{\ell j} * g_{j\ell})$,

for $\ell, k \in \boldsymbol{I}$ with $\ell \neq k$ and $g_{\ell j} * g_{j\ell} < 1$;zero otherwise

4) If $|\boldsymbol{I}| \geq 1$ go to step 1; otherwise stop

Illustrative example

- Consider an oncology trial with k =3 designed to compare a treatment A + SOC versus placebo + SOC for superiority on two primary endpoints PFS and OS.

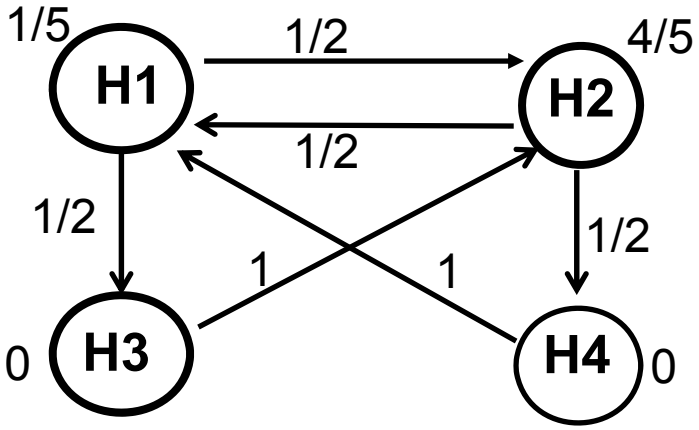
- The trial also has two secondary endpoints SE1 and SE2. The endpoint SE1 can be tested only when the trial is successful on PFS. Similarly, SE2 can be tested only when the trial is successful on OS.

Illustrative example (cont'd)

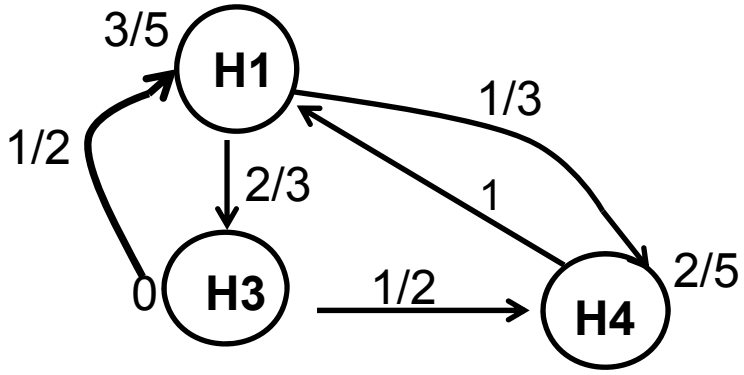
- Therefore, there are 4 hypotheses to test
  - o H_1 and H_2 are primary and are associated with PFS and OS, respectively
  - o H_3 and H_4 are secondary and associated with SE1 and SE2
  - o (H_1, H_3) and (H_2, H_4) are pairs of parent-descendant hypotheses (Maurer et al., 2011)

Graphs (illustrative example)

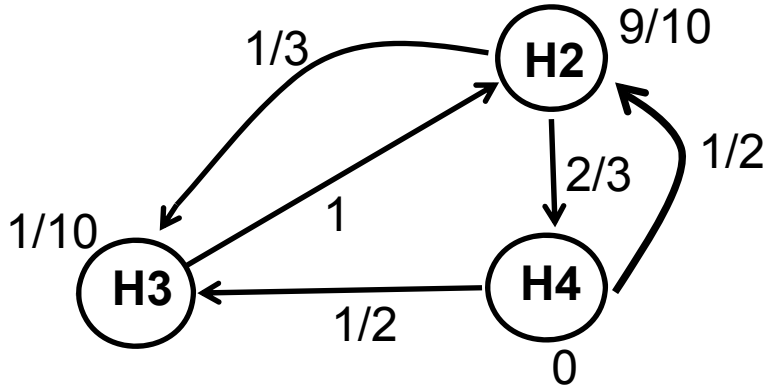
(a) Initial graph



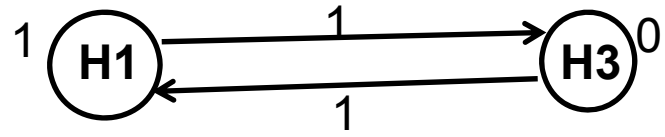
(b) After rejecting H2 in (a)



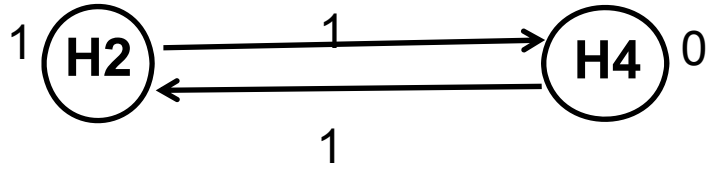
(c) After rejecting H1 in (a)



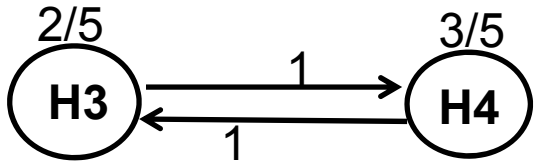
After rejecting H2 and H4



After rejecting H1 and H3



(d) After rejecting H1 and H2



Calculations/decisions at the 1st interim look

- Graph (a): $\boldsymbol{I} = \{1, 2, 3, 4\}$, $(w_i(\boldsymbol{I}), i = 1, ..4) = (1/5, 4/5, 0, 0)$, $\alpha = 0.025$

  Suppose that the O-F type spending function at information times $(t_1, t_2, t_3) = (1/2, 3/4, 1)$

  Recall: O-F type spending function is $f(\gamma, t) = 2[1-\Phi(z_{1-\gamma/2}/t^{1/2})]$

- With the above information, at $t_1 = 0.5$, the alpha critical boundary values are $\{\alpha_{i,t1} (w_i(\boldsymbol{I})\alpha), i = 1, 2, 3, 4\} = \{\alpha_{1,t1} (0.005), \alpha_{2,t1} (0.020), 0, 0\} = \{0.00007, 0.0010, 0, 0\}$ (calculations using software East 6.3)
- Suppose that unadjusted $p_{i,t1} \geq \alpha_{i,t1}$ for all $i = 1, 2, 3, 4$
- Therefore, the trial continues to the 2nd interim look

Calculations/decisions at the 2nd interim look

- Graph (a): $(w_i(\boldsymbol{I}), i = 1, ..4) = (1/5, 4/5, 0, 0)$, but $t_2 = 3/4$.
- Therefore, by the O-F type spending function, at $t_2 = 3/4$, the alpha critical boundary are:

$$\{\alpha_{i,t2}(w_i(\boldsymbol{I})\alpha), i = 1, 2, 3, 4\} = \{\alpha_{1,t2}(0.005), \alpha_{2,t2}(0.020), 0, 0\} = \{0.00117, 0.00690, 0, 0\}$$

- Suppose that, at t_2, $p_{1,t2} = 0.001$, $p_{2,t2} = 0.020$, $p_{3,t2} = 0.040$, $p_{4,t2} = 0.091$.
- Therefore, $H1$ is rejected and one is in graph (c)

Calculations/decisions at the 2nd interim look

- Graph (c): $\boldsymbol{F} = \{2, 3, 4\}$ and $(w_i(\boldsymbol{F}), i = 2, 3, 4) = (9/10, 1/10, 0)$. One can now retest H_2 and also test H_3 by this graph. For this graph:

  $\{\alpha_{i,t2}\,(w_i(\boldsymbol{F})\alpha), i = 2, 3, 4\} = \{\alpha_{2,t2}\,(0.0225), \alpha_{3,t2}\,(0.0025), 0\} = \{0.00802, 0.00047, 0\}$

- Therefore, $H_{\boldsymbol{F}}$ is not rejected, consequently, none of the other hypotheses are rejected, and one proceeds to the 3rd look

Calculations/decisions at the 3rd interim look

- Assume that H1 remain rejected at the 3rd look. Therefore one is in graph (c), where $\boldsymbol{F}$ = {2, 3, 4} and ($w_i(\boldsymbol{F})$, i = 2, 3, 4) = (9/10, 1/10, 0), but t_3 = 1.

- Therefore, again by the O-F type spending function, at t_3 = 1, the alpha critical boundary values are {$\alpha_{i,t3}$ ($w_i(\boldsymbol{F})\alpha$), i = 2, 3, 4} = {$\alpha_{2,t3}$ (0.0225), $\alpha_{3,t3}$ (0.0025), 0} = {0.01988, 0.00234, 0}

- Suppose that, at t_3, $p_{2,t3}$= 0.012, $p_{3,t3}$ = 0.008, $p_{4,t3}$ = 0.041. Therefore, H_2 is rejected and one is in graph (d)

Calculations/decisions at the 3rd interim look

- Graph (d): $\boldsymbol{F}$ = {3, 4} and ($w_i(\boldsymbol{F})$, i = 3, 4) = (2/5, 3/5). Therefore, {$\alpha_{i,t3}$ ($w_i(\boldsymbol{F})\alpha$), i =3, 4} = {$\alpha_{3,t3}$ (0.01), $\alpha_{4,t3}$ (0.015)} = {0.00907, 0.01344}

- Therefore, H_3 is rejected. But H_4 is not rejected because $p_{4,t3}$ = 0.041> $\alpha_{4,t3}$ (0.025) = 0.02200

Concluding Remarks (Part II)

1. Methods based on Bonferroni inequality, as stated, will be rarely used because of low power
2. CTP based procedure of Tang & Geller (1999) simplifies on using weighted Bonferroni tests for intersection hypotheses with weights satisfying consonance property
3. The above approach can be applied for 2 or more hypotheses on using the graphical method with the use of appropriate α-spending functions
4. The above approach leads to α-recycling similar to non-GS procedures but it occurs through the α-spending functions applied

Example of a GS trial design for testing a primary and a secondary endpoint of a trial

2-stage GS trial with 2 endpoints

- Endpoints: X= primary, Y = secondary
- Null hypotheses: H_i: δ_i =0 (i =1, 2) of no treatment effects on X and Y, respectively, tested against 1-sided alternatives
- (X_1, Y_1) and (X_2, Y_2) are pairs of normal Z-test statistics on X and Y, at information times t_1 and t_2 =1, respectively.
- H_2 is tested only after the procedure rejects H_1

2-stage GS trial with 2 endpoints (cont'd)

- *Assumption:* X and Y jointly follow bivariate normal distribution with correlation coefficient of $\rho \geq 0$,

  (c_1, c_2) = boundary values for rejecting H_1;

  (d_1, d_2) = boundary values for rejecting H_2

- Unethical to continue the trial if it is successful in rejecting H_1

Procedure (2-statage design)

Ref: Tamhane et al. (Biometrics 2010)

- Step 1:

  $X_1 \le c_1 \rightarrow$ Go to Step 2

  $X_1 > c_1 \rightarrow$ Reject H_1 and test H_2

  $Y_1 > d_1 \rightarrow$ Reject H_2; else retain it.

  (in either case terminate the trial)

- Step 2:

  $X_2 \le c_2 \rightarrow$ terminate the trial w/o any rejection

  $X_2 > c_2 \rightarrow$ Reject H_1 and test H_2

  $Y_2 > d_2 \rightarrow$ Reject H_2; else retain it.

Determination of boundary values (c_1, c_2) and (d_1, d_2)

- Use of CTP requires considering hypotheses $H_1 \cap H_2$, H_1, and H_2
- For this design, rejecting H_1 at level α also rejects $H_1 \cap H_2$ at the same level α.

  Proof: Consider R_1 and R_2 as rejection regions for H_1, and H_2, respectively.

  $H_1 \cap H_2$ is rejected at level α if Pr (R_1 U R_2} ≤ α. But Pr (R_1 U R_2} = Pr (R_1} ≤ α, as R_2 = is a subset of R_1, because H_2 is tested only after H_1 is rejected.

Determination of boundary values (c_1, c_2) and (d_1, d_2)

- The boundary values (c_1, c_2) for this design can be obtained from the equations:

$$\Pr(X_1 > c_1 \mid H_1) = f_1(\alpha, t_1)$$

$$f_1(\alpha, t_1) + \Pr(X_1 \le c_1 \cap X_2 > c_2 \mid H_1) = f_1(\alpha, t_2 = 1)$$

where, $f_1(\alpha, t)$ is the spending function for the endpoint X

- The boundary values (d_1, d_2) for Y (after rejecting H_1 which rejects $H_1 \cap H_2$) is at level α by the CTP. These boundary values can be obtained using the spending function $f_2(\alpha, t)$ for Y which could be the same as $f_1(\alpha, t_1)$ or different from it.

Example

- Values of c_1 and c_2 , for k =2, $f_1(\alpha, t) = 2[1-\Phi(z_{1-\alpha/2}/t^{1/2})]$, α=0.025, and t_1 = 0.50:

  c_1 = 2.95901, c_2 = 1.96869 (on z-scale)

  $\alpha_{1(x),t1}$ = 0.00154, $\alpha_{2(x),t2}$ = 0.02449 (p-value scale)

- Values of $d = d_1 = d_2$, for k =2 by the PK boundary at α=0.025, and t_1 = 0.50:

  d = 2.17828 (on z-scale), $\alpha_{(y),t1}$ = 0.01469 (p-value scale)

- **<u>Issue</u>: Is it possible to take advantage of the correlation ρ in [0,1) and find $d^* \leq d$ that give FWER control at level α?**

Finding $d^* \leq d$ on taking advantage of correlation between endpoints X *and Y* (cont'd)

- There are 3 null hypotheses configurations: $H_1 \cap H_2$, $H_1 \cap K_2$ and $K_1 \cap H_2$, where K_1 and K_2 are alternatives to H_1 and H_2, respectively.
- The type I error for the first two configurations is ≤ α regardless of the truth and falsity of H_2
- Therefore, such a d^* needs to be found on considering $K_1 \cap H_2$ and the equation:

$$\Pr\{X_1 > c_1 \cap Y_1 > z_y \mid K_1 \cap H_2\} +$$

$$\Pr\{X_1 \leq c_1 \cap X_2 > c_2 \cap Y_2 > z_y \mid K_1 \cap H_2\} = \alpha \qquad (1)$$

) Finding $d^* \leq d$ on taking advantage of correlation between endpoints X and Y (cont'd

- Further, we know that

  $\text{Cov}\{X_1, X_2) = (t_1)^{1/2}$, $\text{Cov}\{X_1, Y_1) = \text{Cov}\{X_2, Y_2) = \rho$, and $\text{Cov}\{X_1, Y_2) = \rho(t_1)^{1/2}$.

- Also, $E(X_1) = \theta(t_1)^{1/2}$, $E(X_2) = \theta$ and $E(Y_j) = 0$ for j =1, 2 (because of $K_1 \cap H_2$ and θ is the drift parameter)

- Further, one can show that, given $X_2 = x_2$, statistics X_1 and Y_2 are independently normally distributed as:

  X_1 is $\mathcal{N}\{ x_2(t_1)^{1/2}, 1\text{-}t_1\}$ and Y_2 is $\mathcal{N}\{ (x_2 - \theta) \rho , 1\text{-} \rho^2\}$

- Therefore, eq. (1) can be written as: (next slide)

Evaluation of Eq. (1)

$$\alpha = 1 - \Phi(c_1 - 6\sqrt{i_1}) - \Phi(z_j) - \Phi_{12}(c_1 - 6\sqrt{i_1}, z_j; \rho)$$

$$+ \int_{c_2-6}^{x} \Phi\left(\frac{c_1 - 6\sqrt{i_1} - u\sqrt{i_1}}{\sqrt{1-i_1}}\right) \Phi\left(\frac{-z_j - u\rho}{\sqrt{1-\rho^2}}\right) \varphi(u)du \qquad (1)$$

Φ is the cumulative distribution functions for the $N(0,1)$ r.v.
Φ_{12} is cumulative distribution function for the standard bivariate normal with correlation coefficient of ρ

Finding d^*:

Assume that values of ρ, t_1, c_1 and c_2 are given.

Then for each $\theta > 0$, one can find z_y that satisfy eq. (1).

Therefore, one can construct a graph $z_y = f(\theta)$ over the interval $\theta > 0$ that satisfy eq. (1).

This will find d^* = the largest z_y so that the RHS of eq. (1) is $\leq \alpha$ for all $\theta > 0$.

Graph of $z_y = f(\theta)$ over the interval $\theta > 0$ satisfying eq. (1)

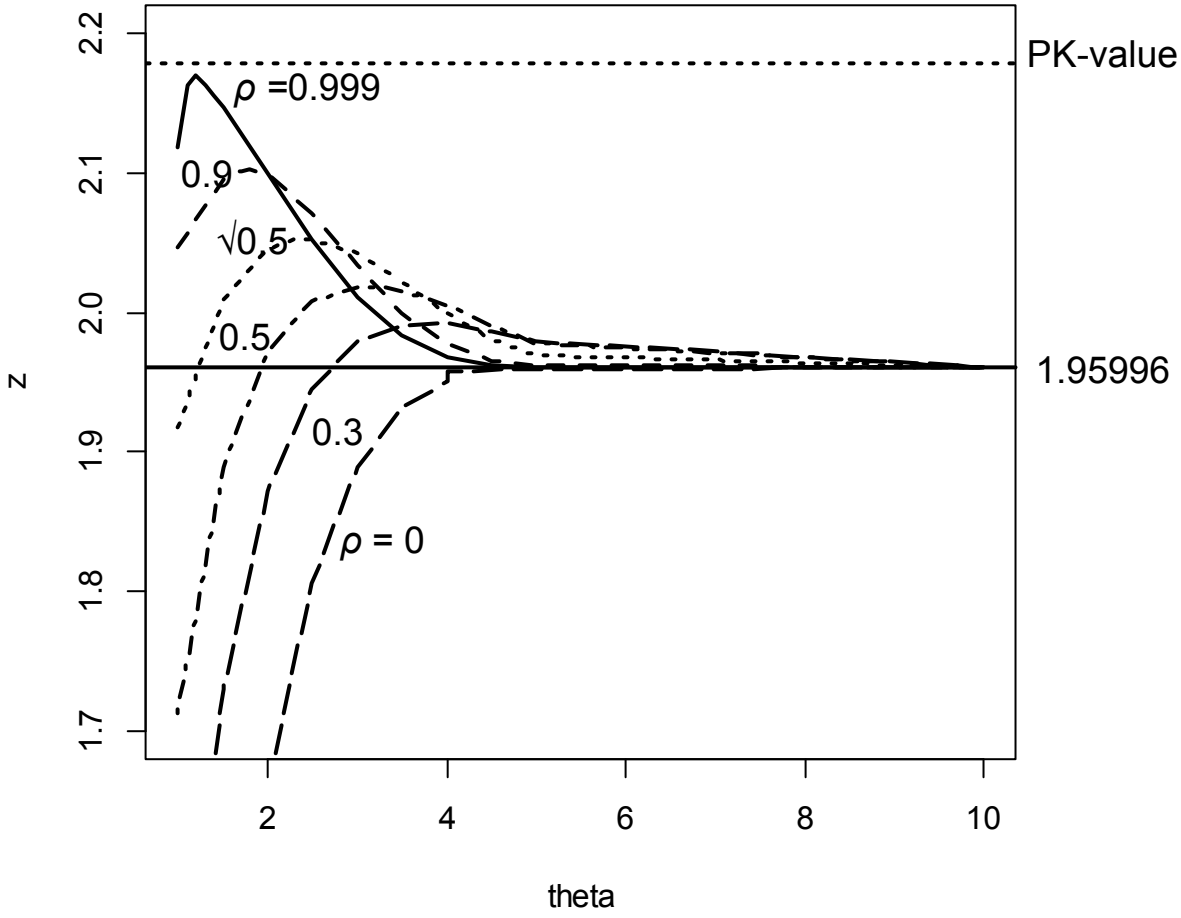


Table: *d**-values for different correlations

$\alpha = 0.025$, $k = 2$, $t_1 = 0.5$, $c_1 = 2.95901$, $c_2 = 1.96869$ (OF-like)

Correlation ρ	***d** (*Z*-scale)**	**α_{d^*} (*p*-value scale)**	**$\theta = \theta^*$**
0.0	1.95996	0.02500	θ^* = all $\theta > 6.5$
0.1	1.96958	0.02444	4.54
0.2	1.98063	0.02382	4.12
0.3	1.99160	0.02321	4.00
0.4	2.00497	0.02248	3.43
0.5	2.01872	0.02176	3.11
0.6	2.03407	0.02097	2.78
√0.5	2.05314	0.02003	2.45
0.8	2.07326	0.01907	2.15
0.9	2.10262	0.01775	1.79
0.99	2.15450	0.01560	1.31
0.999	2.17026	0.01499	1.20
PK value	d =2.17828	α_d= 0.01469	-
Conservative		α/2=0.0125	-

Note: $\theta = \theta^*$ is the value of θ where z_y is maximum on the graph $z_y = f(\theta)$ satisfying eq. (1), for $\theta > 0$.

Concluding Remarks

- In testing a primary and a secondary endpoint null hypotheses for a confirmatory trial, if the correlation between the primary and secondary endpoints cannot be ascertained, then one would test
  1) The primary endpoint null hypothesis by an α-spending function such as OF-like using full alpha
  2) The secondary endpoint can be tested by PK boundary value at the information time the primary endpoint null hypothesis is rejected
- However, if the correlation is known, or if it is known not to exceed ρ_0, then the result of the previous table can be applied for normal distributions of the test statistics

Thank You