COMPARISON BETWEEN TREE GATEKEEPING PROCEDURE AND GRAPHICAL APPROACH IN A PIVOTAL CLINICAL TRIAL WITH MULTIPLE OBJECTIVES AND MULTIPLE ENDPOINTS

07/AUG/2013

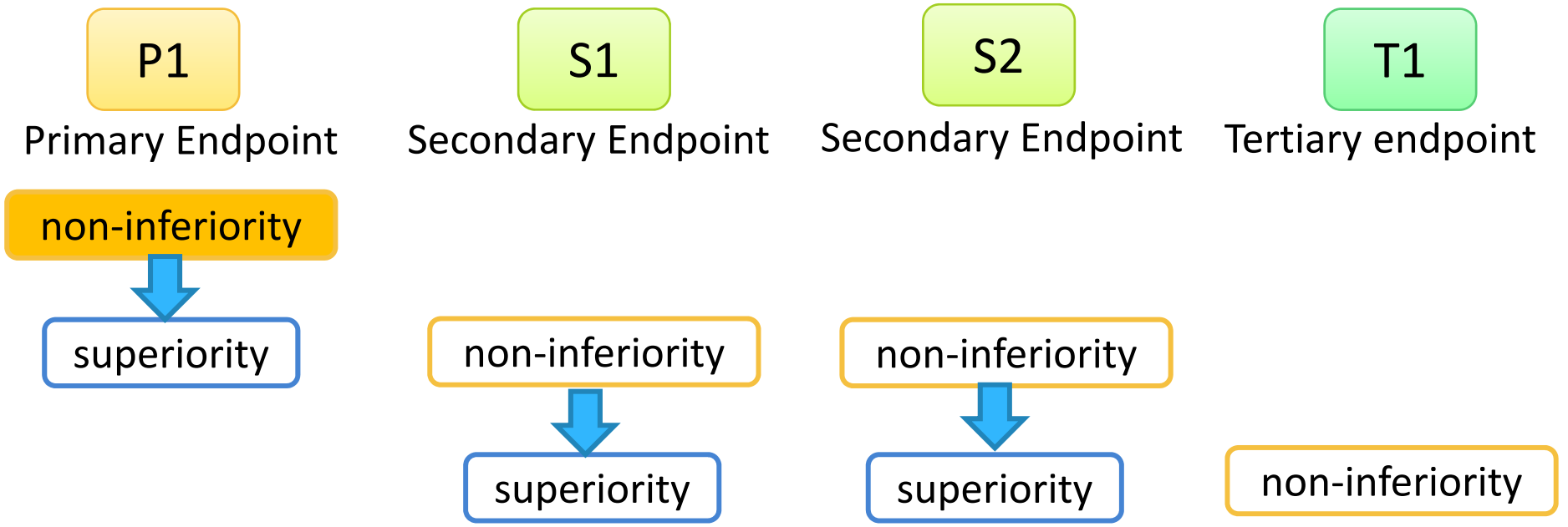
MASAKAZU FUJIWARA, HIDEAKI WATANABE

SHIONOGI & CO., LTD. , JPN

CONTENTS

1. INTRODUCTION

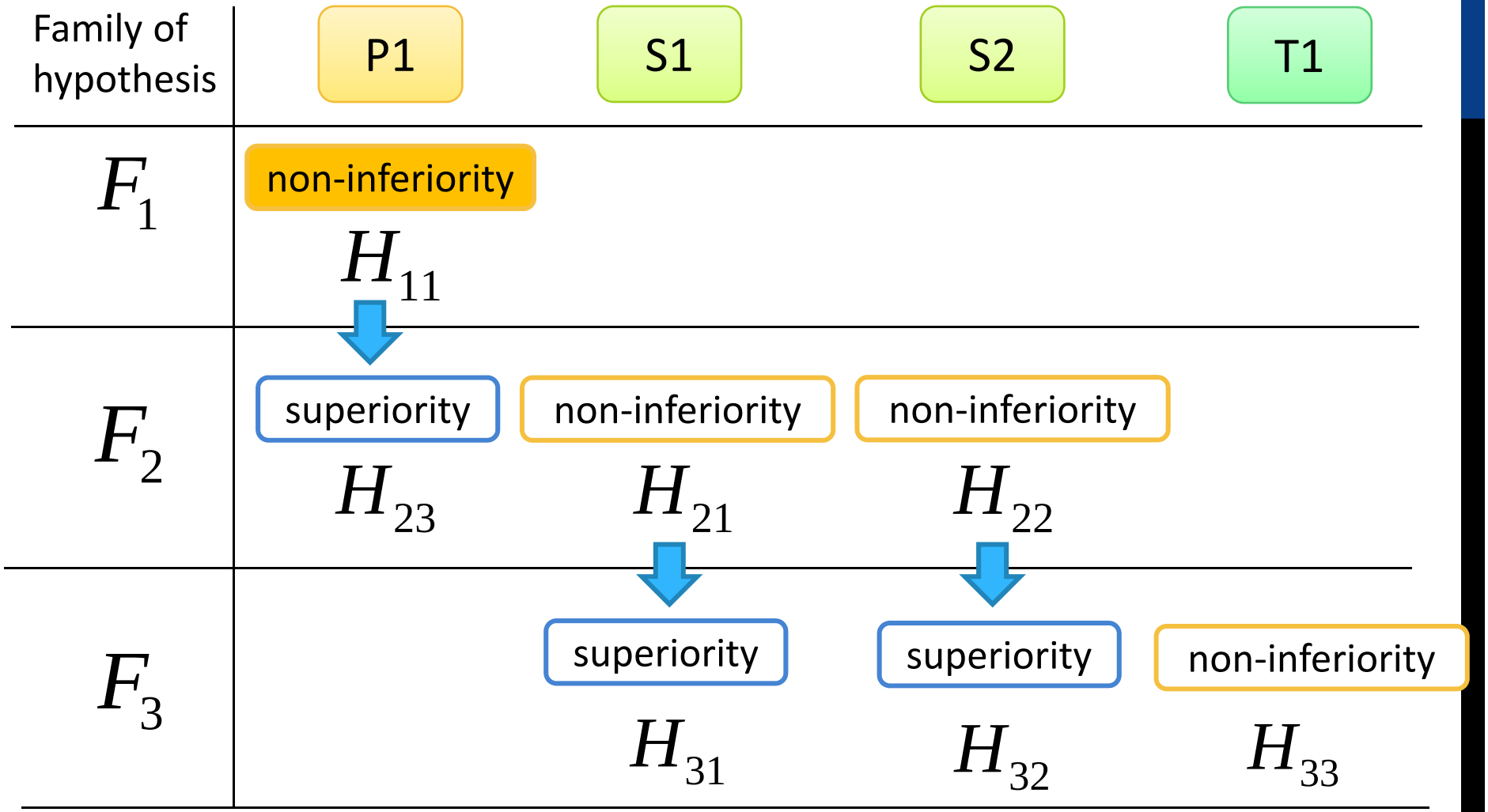
- We assume a pivotal clinical trial for HIV with 2 arms (Drug/Control)



- As to each of primary and secondary endpoints,
  only when non-inferiority can be shown, superiority is tested

When we would like to include these results ***in labeling***, ***FWER (Family-Wise Error Rate) should be controlled***.

SETTING (INDIVIDUAL HYPOTHESIS)



We assume the importance of hypotheses in each family are equal

For convenience, we use these terms in all slides

2. OBJECTIVE

To compare the performance between Bonferroni tree gatekeeping procedure and Graphical approach for assumed pivotal study

Bonferroni tree gatekeeping procedure (BTG)

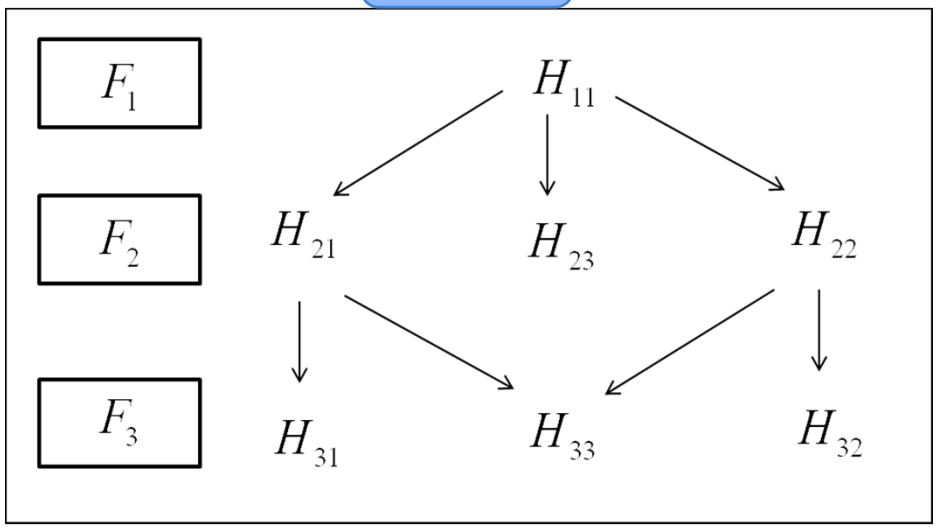
(Dmitrienko, 2007)

Graphical approach (GA)

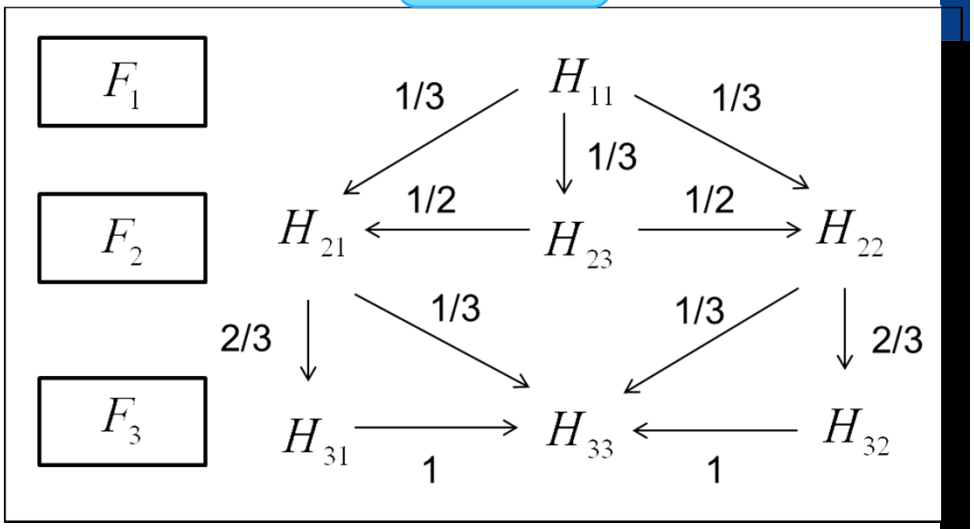
(Bretz, 2009)

OVERVIEW OF BTG AND GA

BTG



GA



Elements influencing on each procedure's performance

1. Rejection set

2. Weights for each individual hypothesis

1. Transition matrix

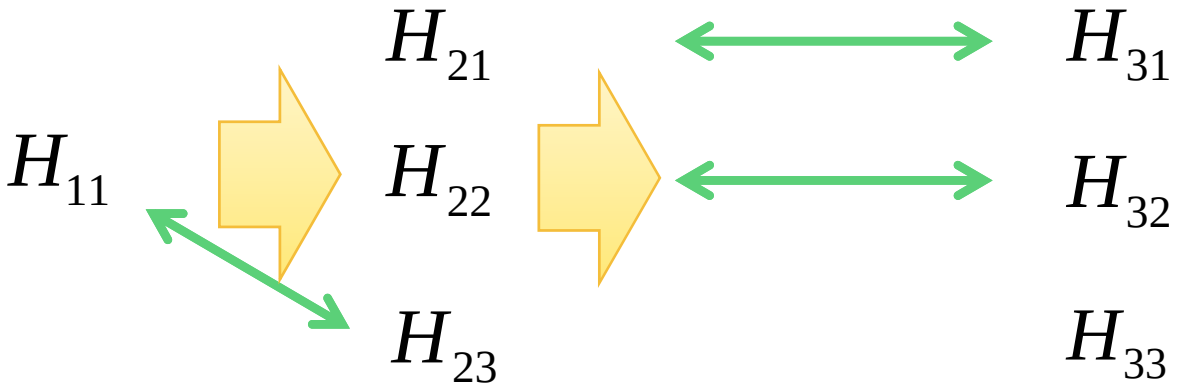
2. Default Alpha α_{ij} for each individual hypothesis

There may be a difference of performance between BTG and GA

WHY DID WE CHOOSE BTG AND GA IN THIS PIVOTAL STUDY ?

- **<u>Two reasons</u>**

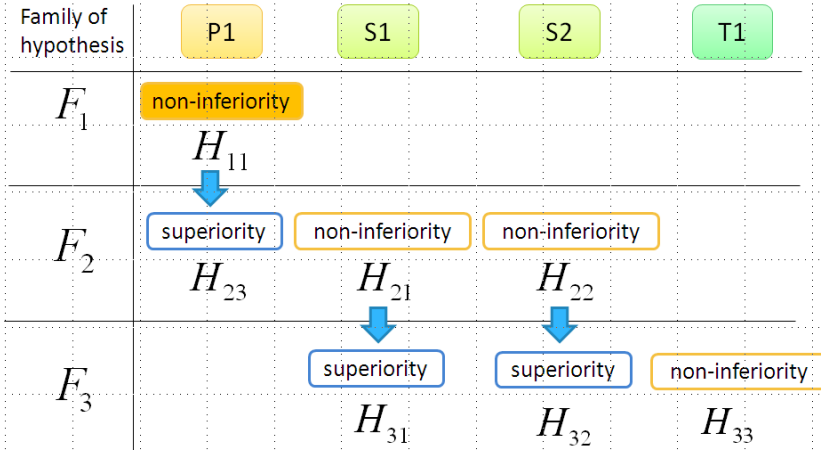
1. **Hierarchical structure between hypotheses**
2. **Test scheme**
   - Only if the non-inferiority for the control drug is verified for P1, S1 and S2, then we can test superiority for these endpoints
   - So, we need the framework that ***<u>only if specific hypothesis in a hypothesis family is rejected, then we can test the particular hypothesis in the next hypothesis family</u>***



3. ASSUMPTION OF SIMULATIONS

Case	Effect size for each hypothesis , Power		Correlations among endpoints
1-1	H_{11} H_{21} H_{22}	0.265 90%	Not considered
	H_{23} H_{31} H_{32} H_{33}	0.203 70%	
1-2	Same as 1-1		Considered
2-1	H_{11}	0.265 90%	Not considered
	H_{21} H_{22} H_{23} H_{33}	0.203 70%	
	H_{31} H_{32}	0.160 50%	
2-2	Same as 2-1		Considered

2arms, Total Sample Size : 600

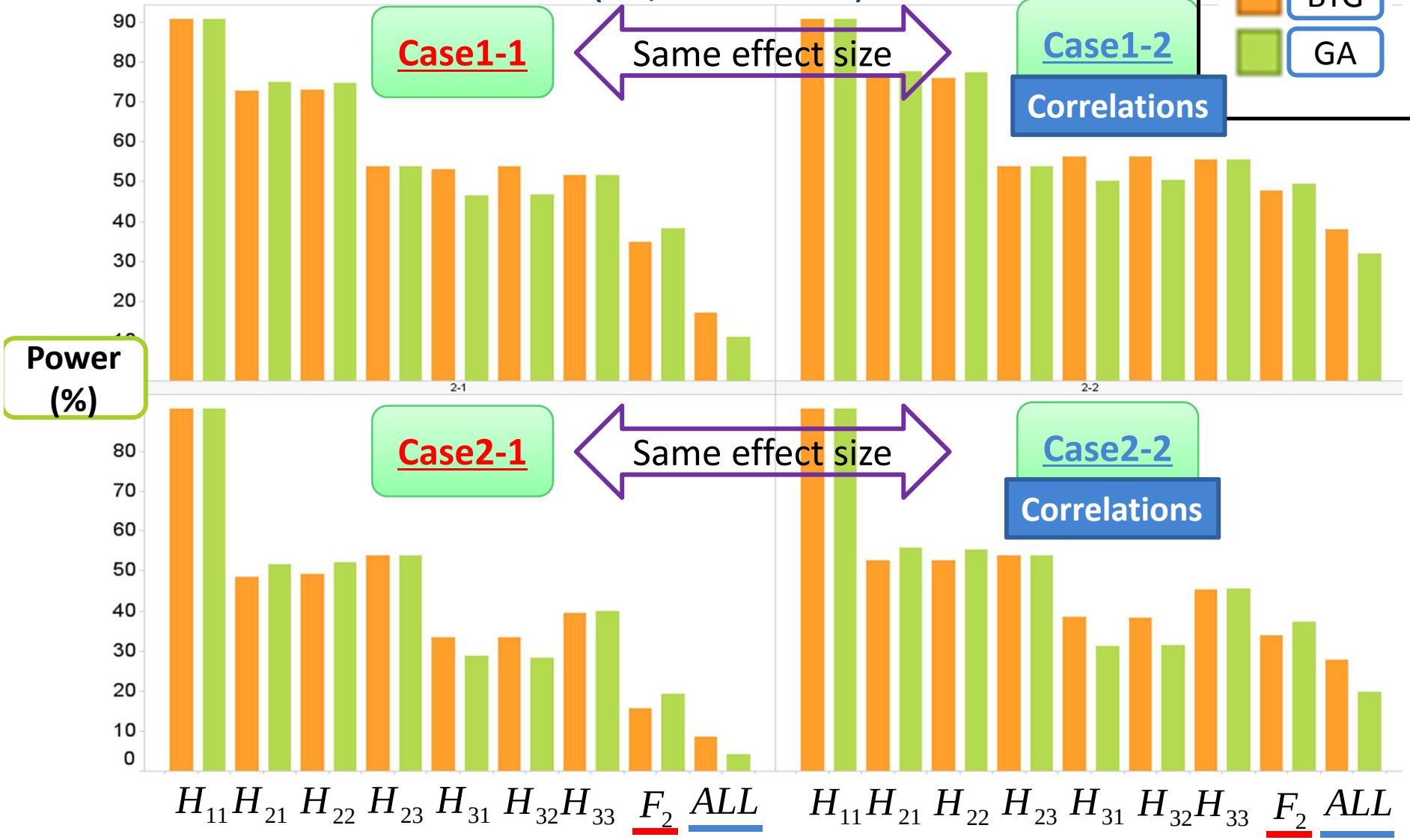


Correlation coefficient matrix

$$(\rho) = \begin{pmatrix} * & 0.6 & 0.6 & 0.6 \\ * & * & 0.8 & 0.6 \\ * & * & * & 0.6 \\ * & * & * & * \end{pmatrix}$$

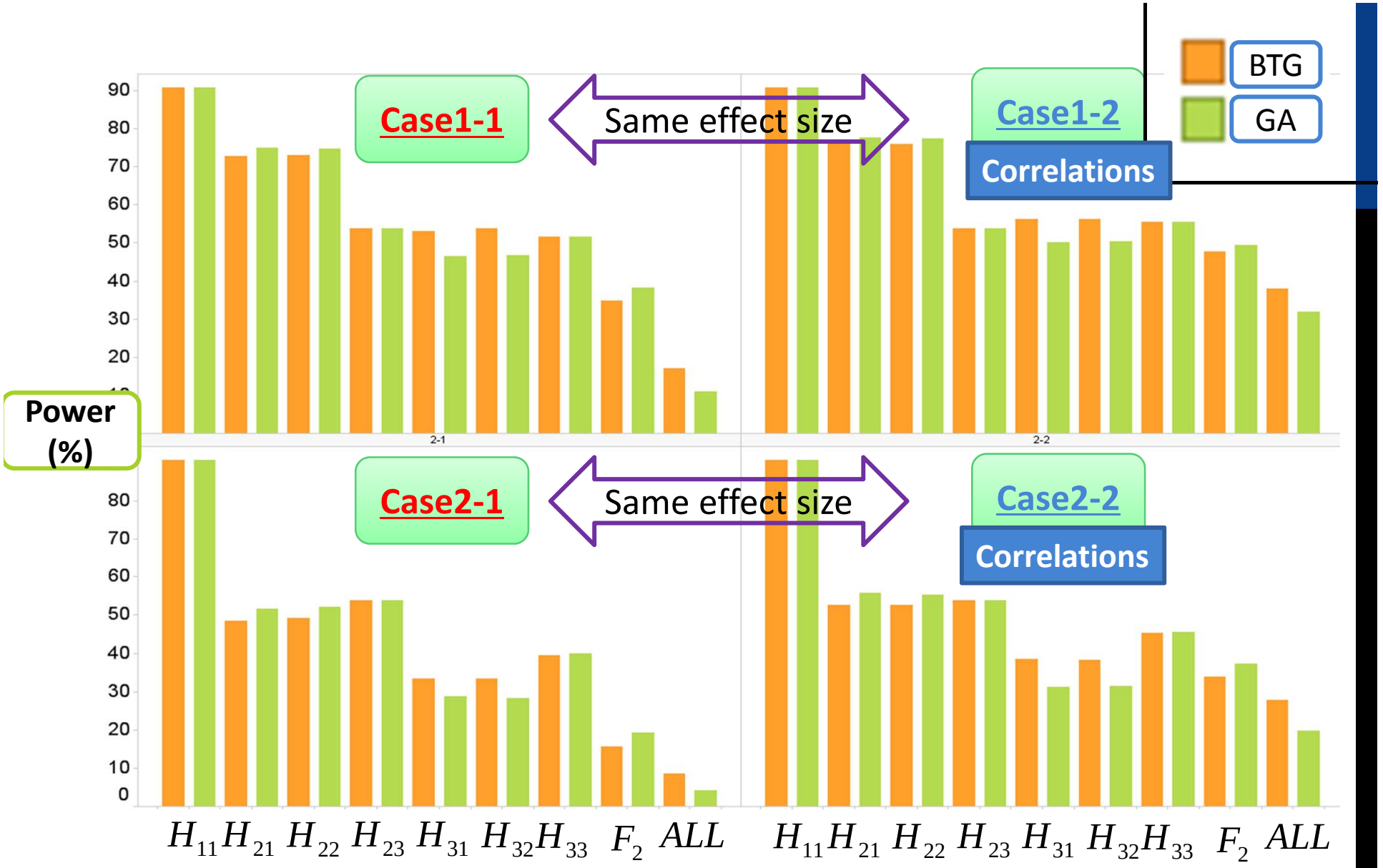
P1 S1 S2 T1

THE RESULT OF SIMULATIONS (10,000 TIMES)

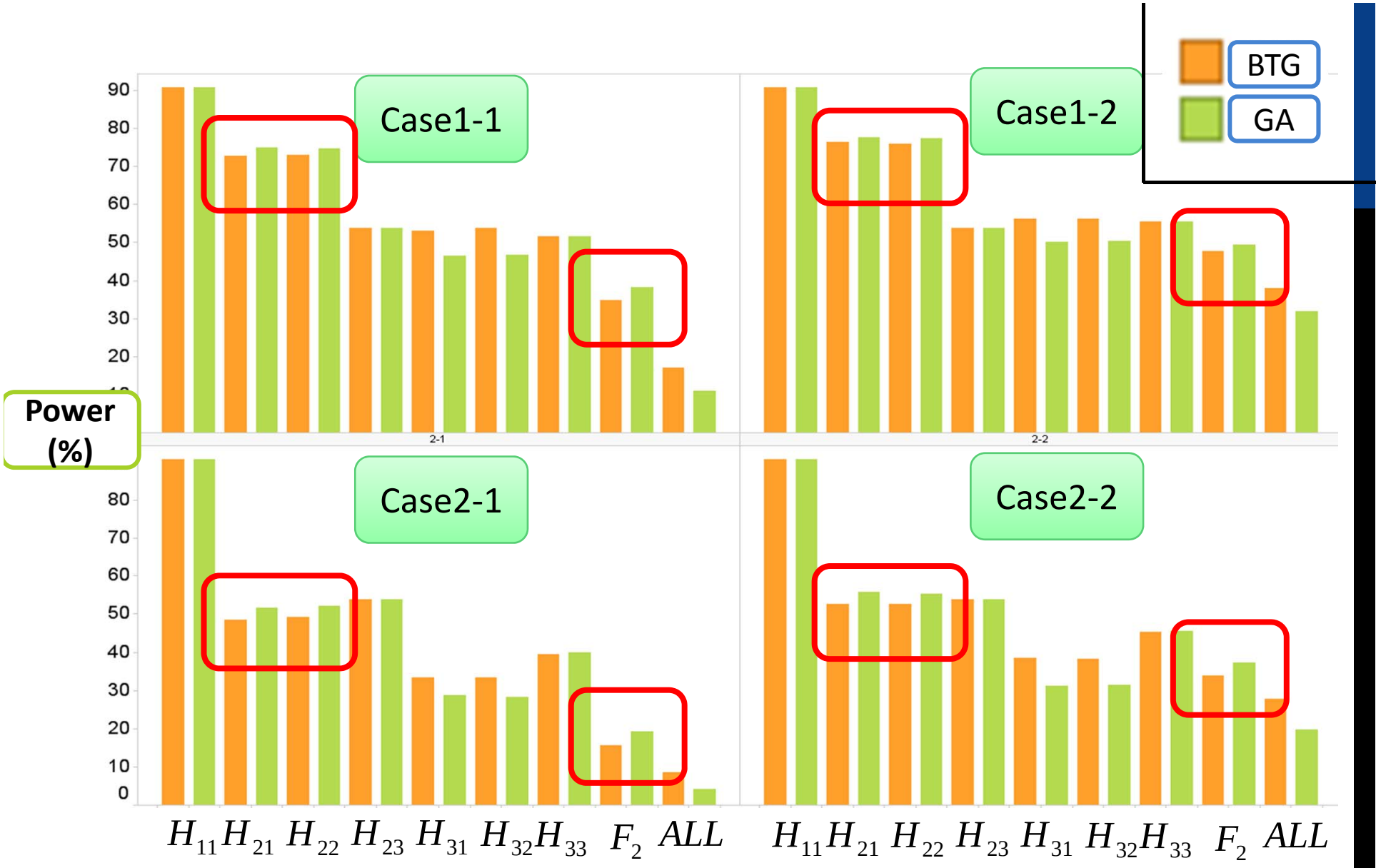


F_2 : the power that ***all hypotheses in*** F_2 were rejected.

ALL : the power that ***all hypotheses*** were rejected.



For both methods, the powers in considering correlations were higher than those in **not considering**, because the power for primary endpoint was 90%

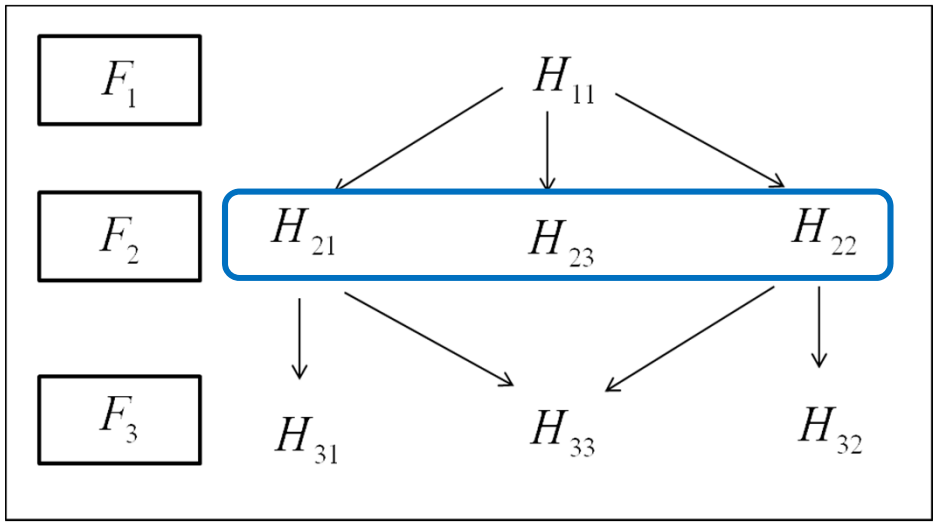


In all cases, the powers of GA for rejecting H_{21}, H_{22} and all hypotheses in F_2 were slightly higher than those of BTG

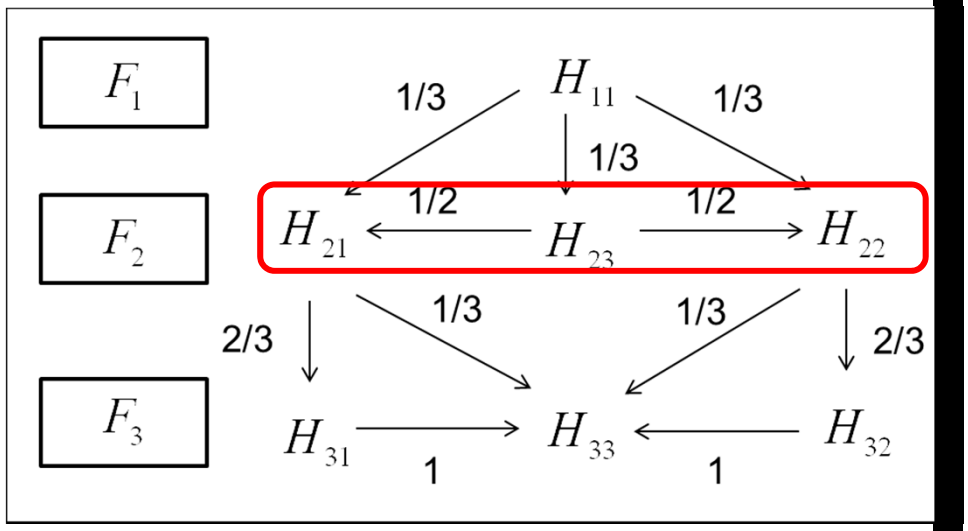
FOCUSED ON F_2

H_{23} was more easily rejected than H_{21}, H_{22} under a given significance level α···

BTG



GA



Adjusted p-values for H_{21}, H_{22}

$3p_{21}, 3p_{22}$

$2p_{21}, 2p_{22}$

If H_{23} was rejected, then GA allocated the significant level efficiently for H_{21}, H_{22} within F_2

In all cases, the powers of BTG for H_{31}, H_{32} and "ALL" were higher than those of GA

FOCUSED ON H_{31}

BTG

$p(H)$: Bonferroni adjusted p-value for each intersection hypotheses H

$\tilde{p}_{ij}$: adjusted p-value for each individual hypotheses H_{ij}

Intersection hypotheses	$v_{31}(H)$
$H_{31,32,33,22,23}$ $H_{31,32,22,23}$	1/6
$H_{31,32,33,23}$	2/9
$H_{31,32,33,22}$ $H_{31,33,22,23}$ $H_{31,33,22}$ $H_{31,33,23}$ $H_{31,32,33}$ $H_{31,32,23}$ $H_{31,22,23}$	1/3
$H_{31,32}$ $H_{31,33}$	1/2
$H_{31,32,22}$ $H_{31,22}$ $H_{31,23}$	2/3
H_{31}	1

If all of $H_{32}, H_{33}, H_{22}, H_{23}$ were not rejected,

$\Rightarrow p(H) = \min_{i,j}\{p_{31} / v_{31}(H)\} = 6p_{31}$ $\Rightarrow \tilde{p}_{31} = \max_H p(H) = 6p_{31}$

$\Rightarrow$ If some of $H_{32}, H_{33}, H_{22}, H_{23}$ were rejected, $\tilde{p}_{31}$ was getting lower

$\Rightarrow$ Minimum $\tilde{p}_{31} = p_{31}$ $\Rightarrow$ α-exhaustive in F_3

FOCUSED ON H_{31}

GA

Intersection hypotheses	$v_{31}(H)$
$H_{31,32,33,22,23}$ $H_{31,32,22,23}$ $H_{31,32,33,23}$ $H_{31,33,22,23}$ $H_{31,33,23}$ $H_{31,22,23}$ $H_{31,23}$	2/9
$H_{31,32,33}$ $H_{31,32,23}$ $H_{31,33,22}$ $H_{31,32,33,22}$ $H_{31,32}$ $H_{31,33}$ $H_{31,22}$ $H_{31,32,22}$ H_{31}	1/3

If all of $H_{32}, H_{33}, H_{22}, H_{23}$ were not rejected, $\tilde{p}_{31} = 4.5\, p_{31}$

Even if $H_{32}, H_{33}, H_{22}, H_{23}$ could be rejected, $\tilde{p}_{31}$ was minimum $3p_{31}$

BTG showed a tendency to get lower adjusted p-values in F_3 than GA, because BTG had α-exhaustive property in F_3

SUMMARY OF SIMULATIONS

In this simulation study,

- Powers of GA to reject the all hypotheses in F_2 were slightly higher than those of BTG
  - GA used the significant level efficiently in the family F_2
- Powers of BTG to reject all hypotheses were higher than those of GA
  - BTG had α-exhaustive property in the last family F_3

4. CONCLUSION

Consideration from the simulation results

- If it is preferable to reject the hypotheses in prior families, we think that GA is appropriate
  - GA can flow the significant level flexibly in prior families
- On the other hand, if it is preferable to reject more hypotheses, we think that BTG is appropriate
  - BTG has α-exhaustive property

Future research

- GA can be modified to have α-exhaustive property
- BTG and GA are Bonferroni-based methods. So, these methods can be expanded to those based on other procedures (ex. Holm, Hochberg etc). Performance of the expanded BTG and GA will be evaluated.

REFERENCES

- Dmitrienko A, Wiens BL, Tamhane AC, Wang X. Tree-structured gatekeeping tests in clinical trials with hierarchically ordered multiple objectives. Statistics in Medicine 2007; 26:2465–2478.
- Hommel G, Bretz F, Maurer W. Powerful short-cuts for multiple testing procedures with special reference to gatekeeping strategies. Statistics in Medicine 2007; 26:4063–4073.
- Dmitrienko A, Wiens BL, Tamhane AC. A note on tree gatekeeping procedures in clinical trials. Statistics in Medicine 2008; 27:3446–3451.
- Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. Statistics in Medicine 2009; 28:586--604.
- Dmitrienko A, Tamhane A, Wang X, Chen X. Stepwise gatekeeping procedures in clinical trial applications. Biometrical Journal 2006; 48:984--991.
- Dmitrienko A, Offen WW, Westfall PH. Gatekeeping strategies for clinical trials that do not require all primary effects to be significant. Statistics in Medicine 2003; 22:2387--2400.
- Wiens BL, Dmitrienko A. The fallback procedure for evaluating a single family of hypotheses. Journal of Biopharmaceutical Statistics 2005; 15:929--942.

world heritage

THANK YOU FOR YOUR ATTENTION !!!