Control of Directional Errors in Fixed Sequence Multiple Testing

Anjana Grandhi
Department of Mathematical Sciences
New Jersey Institute of Technology
Newark, NJ 07102-1982

Wenge Guo
Department of Mathematical Sciences
New Jersey Institute of Technology
Newark, NJ 07102-1982

Joseph P. Romano
Departments of Statistics and Economics
Stanford University
Stanford, CA 94305-4065

Abstract

In this paper, we consider the problem of simultaneously testing many two-sided hypotheses when rejections of null hypotheses are accompanied by claims of the direction of the alternative. The fundamental goal is to construct methods that control the mixed directional familywise error rate, which is the probability of making any type 1 or type 3 (directional)

error. In particular, attention is focused on cases where the hypotheses are ordered as $H_1, \dots, H_n$, so that H_{i+1} is tested only if $H_1, \dots, H_i$ have all been previously rejected. In this situation, one can control the usual familywise error rate under arbitrary dependence by the basic procedure which tests each hypothesis at level α, and no other multiplicity adjustment is needed. However, we show that this is far too liberal if one also accounts for directional errors. But, by imposing certain dependence assumptions on the test statistics, one can retain the basic procedure.

1 Introduction

Directional errors or type 3 errors occur in testing situations with two-sided alternatives when rejections are accompanied by additional directional claims. For example, when testing a null hypothesis $\theta = 0$ against $\theta \neq 0$, rejection of the null hypothesis is often augmented with the decision of whether $\theta > 0$ or $\theta < 0$. In the case of testing a single hypothesis, type 3 error is generally controlled at level α when type 1 error is controlled at level α (and sometimes type 3 error is controlled at level $\alpha/2$). However, in the case of simultaneously testing multiple hypotheses, it is often not known whether additional directional decisions can be made without losing control of the mixed directional familywise error rate (mdFWER), the probability of at least one type 1 or type 3 error. Some methods have been developed in the literature by augmenting additional directional decisions to the existing p-value based stepwise procedures. Shaffer (1980) showed that Holm's procedure (Holm, 1979), augmented with decisions on direction based on the values of test statistics, can strongly control mdFWER under the assumption that the test statistics are independent and under specified conditions on the marginal distributions of the test statistics, but she also showed that counterexamples exist even with two hypotheses. Finner (1994) and Liu (1997) independently proved the same result for the Hochberg procedure (Hochberg, 1988). Finner (1999) generalized the result of Shaffer (1980) to a large class of stepwise or closed multiple test procedures under the same assumptions. Some recent results have been obtained in Guo and Romano (2015).

Several situations occur in practice where hypotheses are ordered in advance, based on relative importance by some prior knowledge (for example in dose-response study, hypotheses of higher dose vs. a placebo are tested before those of lower dose vs. placebo), or there exists a natural hierarchy in tested hypotheses (for example in a clinical trial, secondary endpoints are tested only when the associated primary endpoints are significant), and so on. In such fixed sequence multiple testing situations, it is also desired to make further directional decisions once significant differences are observed. For example, in dose response studies, once the hypothesis of no difference between a dose and placebo is rejected, it is of interest to decide whether the new treatment dose is more or less effective than the placebo. In such cases, the possibility of making type 3 errors must be taken into account.

For control of the usual familywise error rate (FWER) (which does not account for the possibility of additional type 3 errors), the conventional *fixed sequence* multiple testing procedure that strongly controls the FWER under arbitrary dependence, is known to be a powerful procedure in testing situations with pre-ordered hypotheses (Maurer et al., 1995; Wiens, 2003; Wiens and Dmitrienko, 2005). For reviews on recent relevant developments of fixed sequence multiple testing procedures, see Dmitrienko, Tamhane and Bretz (2009) and Dmitrienko, Agostino and Huque (2013). Indeed, suppose null hypotheses $H_1, \ldots, H_n$ are pre-ordered, so that H_{i+1} is tested only if $H_1, \ldots, H_i$ have all been rejected. The probability mechanism generating the data is P and H_i asserts that $P \in \omega_i$, some family of data generating distributions. In such case, it is easy to see that each H_i can be tested at level α in order to control the FWER at level α, so that no adjustment for multiplicity is required. The argument is simple and goes as follows. Fix any given P such that at least one H_i is true (or otherwise the FWER is 0 anyway). If H_1 is true, i.e. $P \in \omega_1$, then a type 1 error occurs if and only if H_1 is rejected, and so the FWER is just the probability H_1 is rejected, which is assumed controlled at level α when testing H_1. If H_1 is false, just let f be the smallest index corresponding to a true null hypothesis, i.e. H_f is true but $H_1, \ldots, H_{f-1}$ are all false. In this case, a type 1 error occurs if and only if H_f is rejected, which is assumed to be controlled at level α.

In fact, in situations where ordering is not specified, the above result suggests it may be worthwhile to think about hypotheses in order of importance so that potentially false hypotheses are more easily detected. Indeed, as is well-known, when the number n of tested hypotheses is large, control of the FWER is often so stringent that often no rejections can be detected, largely due to the multiplicity of tests and the need to find significance at very low levels (as required, for example, in the Bonferroni method with n large). On the other hand, under a specified ordering, each test is carried out at the same conventional level.

To our knowledge, no one explores the possibility of making additional directional decisions for such fixed sequence procedures. In this paper, we introduce such fixed sequence procedures augmented with additional directional decisions and discuss its mdFWER control under independence and some dependence. For such directional procedures, its simple fixed sequence structure of the tested hypotheses makes the notoriously challenging problem of controlling the mdFWER under dependence a little easier to handle than stepwise procedures.

Throughout this work, we consider the problem of testing n two-sided hypotheses $H_1, \ldots, H_n$ specified as follows:

$$H_i : \theta_i = 0 \quad \text{vs.} \quad H_i^{'} : \theta_i \neq 0, \quad i = 1, \ldots, n. \tag{1}$$

We assume the hypotheses are ordered in advance, either using some prior knowledge about the importance of the hypotheses or by some other specified criteria, so that H_1 is tested first and H_i is only tested if $H_1, \ldots, H_{i-1}$ are all rejected. We also assume that, for each i, a test statistic T_i and p-value P_i are available to test H_i (as a single test). For a rejected hypothesis H_i, we decide on the sign of the parameter θ_i by the sign of the corresponding test statistic T_i, i.e., we conclude $\theta_i > 0$ if $T_i > 0$ and vice versa. The errors that might occur while testing these hypotheses are type 1 and type 3 errors. A *type 1 error* occurs when a true H_i is falsely rejected. A *type 3 error* occurs when a false H_i is correctly rejected but the claimed sign of the parameter θ_i is wrong. Then, the mdFWER is the probability of making at least a type 1 or type 3 error, and it is desired that this error rate is no bigger than α for all possible data generating distributions in the model.

We make a few standard assumptions about the test statistics. Let $T_i \sim F_{\theta_i}(\cdot)$ for some continuous cumulative distribution function $F_{\theta_i}(\cdot)$ having parameter θ_i. In general, most of our results also apply through the same arguments when the family of distributions of T_i depends on i, though for simplicity of notation, the notation is suppressed. We assume that F_0 is symmetric about 0 and F_{θ_i} is stochastically increasing in θ_i. Various dependence assumptions between the test statistics will be used throughout the paper. (Some of the results can generalize outside this parametric framework. Of course, for many problems, approximations are used to construct marginal tests and the approximate distributions of the T_i are often normal, in which case our exact finite sample results will hold approximately as well.) Let $c_1 = F_0^{-1}(\alpha/2)$ and $c_2 = F_0^{-1}(1-\alpha/2)$, so that a marginal level α test of H_i rejects if $T_i < c_1$ or $T_i > c_2$. For testing H_i vs. $H_i^{'}$, rejections are based on large values of $|T_i|$ and the corresponding two-sided p-value is defined by

$$P_i = 2\min\{F_0(T_i), 1 - F_0(T_i)\}, \quad i = 1, \ldots, n. \tag{2}$$

We assume that the p-value P_i is distributed as U(0,1) when $\theta_i = 0$.

The rest of the paper is organized as follows. In Section 2, we consider the problem of mdFWER control under no dependence assumptions on the test statistics. Unlike control of the usual FWER where each test can be constructed at level α, it is seen that H_i can only be tested at a much smaller level $\alpha/2^{i-1}$. Such a rapid decrease in the critical values used motivates the study of the problem under various dependence assumptions. In Section 3 we introduce a directional fixed sequence procedure and prove that this procedure controls the mdFWER under independence. In Sections 4 and 5 we further discuss its mdFWER control under positive dependence. Section 6 presents some concluding remarks. All proofs are deferred to Section 7.

2 The mdFWER Control Under Arbitrary Dependence

A general fixed sequence procedure based on marginal p-values must specify the critical level α_i that is used for testing H_i, in order for the resulting procedure to control the md-FWER at level α. When controlling the FWER without regard to type 3 errors, each α_i can be as large as α. However, Theorem 1 below shows that by using the critical constant $\alpha_i = \alpha/2^{i-1}$, the mdFWER is controlled at level α. Moreover, we show that these critical constants are unimprovable. Formally, the optimal procedure is defined as follows.

Procedure 1 (Directional fixed sequence procedure under arbitrary dependence)

- *Step 1: If $P_1 \leq \alpha$ then reject H_1 and continue to test H_2 after making directional decision on θ_1: conclude $\theta_1 > 0$ if $T_1 > 0$ or $\theta_1 < 0$ if $T_1 < 0$. Otherwise, accept all the hypotheses and stop.*

- *Step i: If $P_i \leq \alpha/2^{i-1}$ then reject H_i and continue to test H_{i+1} after making directional decision on θ_i: conclude $\theta_i > 0$ if $T_i > 0$ or $\theta_i < 0$ if $T_i < 0$. Otherwise, accept the remaining hypotheses $H_i, \ldots, H_n$.*

In the following, we discuss the mdFWER control of Procedure 1 under arbitrary dependence of the p-values. When testing a single hypothesis, the mdFWER of Procedure 1 reduces to the type 1 or type 3 error rate depending on whether $\theta = 0$ or $\theta \neq 0$, and Procedure 1 reduces to the usual p-value based method along with the directional decision for the two-sided test. The following lemma covers this case.

Lemma 1 *Consider testing the single hypothesis $H : \theta = 0$ against $H' : \theta \neq 0$ at level α, using the usual p-value based method along with a directional decision. If H is a false null hypothesis, then the type 3 error rate is bounded above by $\alpha/2$.*

Generally, when simultaneously testing n hypotheses, by using Lemma 1 and mathematical induction, we have the following result holds.

Theorem 1 *For Procedure 1 defined as above, the following conclusions hold.*

(i) This procedure strongly controls the mdFWER at level α under arbitrary dependence of the p-values.

(ii) One cannot increase even one of the critical constants $\alpha_i = \alpha/2^{i-1}, i = 1, \ldots, n$, while keeping the remaining fixed without losing control of the mdFWER.

In fact, the proof shows that no strong parametric assumptions are required. However, the rapid decrease in critical values $\alpha/2^{i-1}$ makes rejection of additional hypotheses difficult. Thus, it is of interest to explore how dependence assumptions can be used to increase these critical constants while maintaining control of the mdFWER. The assumptions and methods will be described in the remaining sections.

Remark 1 Instead of Procedure 1, let us consider the conventional fixed sequence procedure with the same critical constant α augmented with additional directional decisions, which is defined in Section 3 as Procedure 2. By using Bonferroni inequality and Lemma 1, we can prove that the mdFWER of this procedure is bounded above by $\frac{n+1}{2}\alpha$. Thus, the modified version of the procedure, which has the same critical constant $\frac{2\alpha}{n+1}$, strongly controls the mdFWER at level α under arbitrary dependence of p-values. However, it is unclear if such critical constant can be further improved without losing the control of the mdFWER.

3 The mdFWER Control Under Independence

We further make the following assumptions on the distribution of the test statistics.

Assumption 1 (Independence) *The test statistics, $T_1, \ldots, T_n$, are mutually independent.*

Of course, it follows that the p-values $P_1, \ldots, P_n$ are mutually independent as well.

As will be seen, it will be necessary to make further assumptions on the family of distributions for each marginal test statistic.

Definition 1 (Monotone Likelihood Ratio (MLR)) *A family of probability density functions $f_\delta(\cdot)$ is said to have monotone likelihood ratio property if, for any two values of the parameter δ, $\delta_2 > \delta_1$ and any two points $x_2 > x_1$,*

$$\frac{f_{\delta_2}(x_2)}{f_{\delta_1}(x_2)} \geq \frac{f_{\delta_2}(x_1)}{f_{\delta_1}(x_1)}, \tag{3}$$

or equivalently,

$$\frac{f_{\delta_1}(x_1)}{f_{\delta_1}(x_2)} \geq \frac{f_{\delta_2}(x_1)}{f_{\delta_2}(x_2)}. \tag{4}$$

Definition 1 means that, for fixed $x_1 < x_2$, the ratio $\frac{f_\delta(x_1)}{f_\delta(x_2)}$ is non-increasing in δ. Two direct implications of Definition 1 in terms of the cdf $F_\delta(\cdot)$ are

$$\frac{F_{\delta_1}(x_2)}{F_{\delta_1}(x_1)} \leq \frac{F_{\delta_2}(x_2)}{F_{\delta_2}(x_1)}, \tag{5}$$

and

$$\frac{1 - F_{\delta_1}(x_2)}{1 - F_{\delta_1}(x_1)} \leq \frac{1 - F_{\delta_2}(x_2)}{1 - F_{\delta_2}(x_1)}. \tag{6}$$

Assumption 2 (MLR Assumption) *The family of marginal distributions of the T_i has monotone likelihood ratio.*

Based on the conventional fixed sequence multiple testing procedure, we define a directional fixed sequence procedure as follows, which is the conventional fixed sequence procedure augmented with directional decisions. In other words, any hypothesis is tested at level α, and as will be seen under the specified conditions, no reduction in critical values is necessary in order to achieve mdFWER control.

Procedure 2 (Directional fixed sequence procedure)

- *Step 1: If $P_1 \leq \alpha$, then reject H_1 and continue to test H_2 after making a directional decision on θ_1: conclude $\theta_1 > 0$ if $T_1 > 0$ or $\theta_1 < 0$ if $T_1 < 0$. Otherwise, accept all the hypotheses and stop.*

- *Step i: If $P_i \leq \alpha$, then reject H_i and continue to test H_{i+1} after making a directional decision on θ_i: conclude $\theta_i > 0$ if $T_i > 0$ or $\theta_i < 0$ if $T_i < 0$. Otherwise, accept the remaining hypotheses, $H_i, \ldots, H_n$.*

For Procedure 2, in the case of $n = 2$, we derive a simple expression for the mdFWER in Lemma 2 below and prove its mdFWER control in Lemma 3 by using such simple expression.

Lemma 2 *Consider testing two hypotheses $H_1 : \theta_1 = 0$ and $H_2 : \theta_2 = 0$, against both sided alternatives, using Procedure 2 at level α. Let $c_1 = F_0^{-1}(\alpha/2)$ and $c_2 = F_0^{-1}(1 -$*

$\alpha/2$). *When* $\theta_2 = 0$*, the following result holds.*

$$mdFWER = \begin{cases} \alpha + F_{\theta_1}(c_1) - F_{\theta_1}(c_2) + F_{(\theta_1,0)}(c_2, c_2) - F_{(\theta_1,0)}(c_2, c_1) & \text{if } \theta_1 > 0 \\ 1 + F_{\theta_1}(c_1) - F_{\theta_1}(c_2) + F_{(\theta_1,0)}(c_1, c_1) - F_{(\theta_1,0)}(c_1, c_2) & \text{if } \theta_1 < 0. \end{cases} \tag{7}$$

In the above, $F_{\theta_1,\theta_2}(\cdot,\cdot)$ *refers to the joint c.d.f. of* (T_1, T_2)*. Then, under Assumption 1 (independence) , (7) can be simplified as*

$$mdFWER = \begin{cases} \alpha + F_{\theta_1}(c_1) - \alpha F_{\theta_1}(c_2) & \text{if } \theta_1 > 0 \\ 1 + \alpha F_{\theta_1}(c_1) - F_{\theta_1}(c_2) & \text{if } \theta_1 < 0. \end{cases} \tag{8}$$

Lemma 3 *Under Assumption 1 (independence) and Assumption 2 (MLR), Procedure 2 strongly controls the mdFWER when* $n = 2$*.*

Generally, for testing any n hypotheses, by using mathematical induction and Lemma 3, we also prove the mdFWER control of Procedure 2 under the same assumptions as in the case of $n = 2$.

Theorem 2 *Under Assumption 1 (independence) and Assumption 2 (MLR), Procedure 2 strongly controls the mdFWER at level* α*.*

Many families of distributions have the MLR property: normal, uniform, logistic, Laplace, Student's t, generalized extreme value, exponential familes of distributions, etc. However, it is also important to know whether or not the above results fail without the MLR assumption. A natural family of distributions to consider without the MLR property is the Cauchy family; indeed, Shaffer (1980) used this family to obtain a counterexample for the directional Holm procedure while testing p-value ordered hypotheses. We now show that Procedure 2 fails to control the mdFWER for this family of distributions with corresponding cdf $F_\theta(x) = 0.5 + \frac{1}{\pi}\arctan(x - \theta)$, even under independence.

Lemma 2 can be used to verify the calculation for the case of $n = 2$ with $\theta_1 > 0$ and $\theta_2 = 0$; specifically, see (8). Indeed, we just need to show show

$$F_{\theta_1}(-c) = F_0(-c - \theta_1) > \alpha F_{\theta_1}(c) = \alpha F_0(c - \theta_1) \ , \tag{9}$$

where c is the $1 - \alpha/2$ quantile of the standard Cauchy distribution, given by $\tan[\pi(1 - \alpha)/2]$. Take $\alpha = 0.05$, so $c = 12.7062$. Then, the above inequality (9) is violated for example by $\theta_1 = 100$. The left side is approximately $F(-112.7) \approx 0.002824$ while the right side is

$$0.05 \times F(-87.3) = 0.05 \times 0.0036 = 0.00018.$$

4 Extension to Positive Dependence

Clearly, the assumption of independence is of limited utility in multiple testing, as many tests are usually carried out on the same data set. Thus, it is important to generalize the results of the previous section to cover some more general cases. As is typical in the multiple testing literature (Benjamini and Yekutieli, 2001; Sarkar, 2002; Sarkar and Guo, 2010, etc), assumptions of positive regression dependence will be used.

Before defining the assumptions, for convenience, we introduce several notations below. Among the prior-ordered hypotheses $H_1, \ldots, H_n$, let i_0 denote the index of the first true null hypothesis, n_1 denote the number of all false nulls, and $T_{i_1}, \ldots, T_{i_{n_1}}$ denote the corresponding false null test statistics. Specifically, if all H_i's are false, let $i_0 = n + 1$.

Assumption 3 *The false null test statistics along with parameters, $\theta_{i_1} T_{i_1}, \ldots, \theta_{i_{n_1}} T_{i_{n_1}}$, are positively regression dependent in the sense of*

$$E\left\{\phi(\theta_{i_1} T_{i_1}, \ldots, \theta_{i_{n_1}} T_{i_{n_1}}) \mid \theta_{i_k} T_{i_k} \geq u\right\} \uparrow u, \tag{10}$$

for each $\theta_{i_k} T_{i_k}$ and any (coordinatewise) non-decreasing function ϕ.

Assumption 4 *The first true null statistic, T_{i_0}, is independent of all false null statistics $T_{i_k}, k = 1, \ldots, n_1$ with $i_k < i_0$.*

Theorem 3 *Under Assumptions 2 - 4, Procedure 2 strongly controls the mdFWER at level α.*

Corollary 1 *When all tested hypotheses are false, Procedure 2 strongly controls the mdFWER at level α under Assumptions 2 - 3.*

Remark 2 In Theorem 3, we note that specifically, when all of the tested hypotheses are false, Assumption 4 is automatically satisfied. Generally, consider the case of any combination of true and false null hypotheses where Assumption 4 is not imposed. Without loss of generality, suppose $\theta_i > 0, i = 1, \ldots, n-1$ and $\theta_n = 0$, that is, the first $n-1$ hypotheses are false and the last one is true. Under Assumptions 2-3, if T_n (or $-T_n$) and $T_1, \ldots, T_{n-1}$ are positively regression dependent, then the mdFWER of Procedure 2 when testing $H_1, \ldots, H_n$ is, for any n, bounded above by

$$
\begin{aligned}
& \text{Pr(make at least one type 3 error when testing } H_1, \ldots, H_{n-1} \text{ or } T_n \notin (c_1, c_2)) \\
\leq \;& \lim_{\theta_n \to 0+} \text{Pr(make at least one type 3 error when testing } H_1, \ldots, H_n) \\
& \quad + \lim_{\theta_n \to 0+} Pr(T_n \geq c_2) \\
\leq \;& \alpha + \alpha/2 = 3\alpha/2.
\end{aligned}
$$

The first inequality follows from the fact that when $\theta_n \to 0+$, H_n can be interpreted as a false null hypothesis with $\theta_n > 0$, and thus one type 3 error is made if H_n is rejected and $T_n \leq c_1$. The second inequality follows from Corollary 1 and Lemma 1.

Based on the above inequality, a modified version of Procedure 2, the directional fixed sequence procedure with the critical constant $2\alpha/3$, strongly controls the mdFWER at level α under Assumptions 2-3 and the above additional assumption.

Remark 3 In the above remark, further, if we do not make any assumption regarding dependence between the true null statistic T_n and the false null statistics $T_1, \ldots, T_{n-1}$. Then, by Theorem 3, the mdFWER of Procedure 2 when testing $H_1, \ldots, H_n$ is bounded above by

$$
\begin{aligned}
& \text{Pr(make at least one type 3 error when testing } H_1, \ldots, H_{n-1}) \\
& \quad + \text{Pr(make type 1 error when testing } H_n) \\
\leq \;& \alpha + \alpha = 2\alpha.
\end{aligned}
$$

Therefore, an alternative modified version of Procedure 2, the directional fixed sequence procedure with the critical constant $\alpha/2$, strongly controls the mdFWER at level α only under Assumptions 2-3.

5 Further Extensions to Positive Dependence

We now develop alternative results to show that Procedure 2 can control mdFWER even under certain dependence between the false null and true null statistics. We relax the assumption of independence that the false null statistics are independent of the first true null statistic, and consider a slightly strong version of the conventional positive regression dependence on subset of true null statistics (PRDS) (Benjamini and Yekutieli, 2001), which is given below.

Assumption 5 *The false null test statistics, $T_1, \dots, T_{i_0-1}$ and the first true null statistic T_{i_0}, are positive regression dependent in the sense of*

$$E\left\{\phi(T_1, \dots, T_{i_0-1}) \mid T_{i_0} \geq u, T_1, \dots, T_j\right\} \uparrow u, \tag{11}$$

for any given $j = 1, \dots, i_0 - 1$, any given values of $T_1, \dots, T_j$ and any (coordinatewise) non-decreasing function ϕ.

We firstly consider the case of $n = 2$, that is, while testing two hypotheses, and show control of the mdFWER of Procedure 2 when the test statistics are positively regression dependent in the sense of Assumption 5.

Proposition 1 *Under Assumptions 2 and 5, the mdFWER of Procedure 2 is strongly controlled at level α when $n = 2$.*

Specifically, in the case of bivariate normal distribution, Assumption 2 is satisfied and two test statistics T_1 and T_2 are always positively or negatively regression dependent. As in the proof of Proposition 1, to show the mdFWER control of Procedure 2, we only need to consider the case of $\theta_1 \neq 0$ and $\theta_2 = 0$. Thus, if T_1 and T_2 are negatively regression dependent, we can choose $-T_2$ as the statistic for testing H_2 and Assumption 5 is still satisfied. By Proposition 1, we have the following corollary holds.

Corollary 2 *Under the case of bivariate normal distribution, the mdFWER of Procedure 2 is strongly controlled at level α when $n = 2$.*

We now consider the case of three hypotheses. The general case will ultimately be considered, but is instructive to discuss the case separately due to the added multivariate MLR condition, which is described as follows.

Let $f(x|T_1)$ and $g(x|T_1)$ denote the probability density functions of T_2 and T_3 conditional on T_1, respectively.

Assumption 6 (Bivariate Monotone Likelihood Ratio (BMLR)) *For any given value of T_1, $f(x|T_1)$ and $g(x|T_1)$ have the monotone likelihood ratio (MLR) property in x, i.e., for any $x_2 > x_1$, we have*

$$\frac{f(x_2|T_1)}{g(x_2|T_1)} \geq \frac{f(x_1|T_1)}{g(x_1|T_1)}. \tag{12}$$

Proposition 2 *Under Assumptions 2, 3, 5, and 6, the mdFWER of Procedure 2 is strongly controlled at level α when $n = 3$.*

Remark 4 In the case of three hypotheses, suppose that the test statistics $T_i, i = 1, \ldots, 3$ are trivariate normally distributed with the mean θ_i. Without loss of generality, assume $\theta_i > 0, i = 1, 2$ and $\theta_3 = 0$, that is, H_1 and H_2 are false and H_3 is true. Let $\Sigma = (\sigma_{ij}), i, j = 1, \ldots, 3$, denote the variance-covariance matrix of T_i's. It is easy to see that Assumption 2 is always satisfied. Also, when $\sigma_{ij} \geq 0$ for $i \neq j$, Assumption 3 and Assumption 5 are satisfied. Finally, when $\sigma_{22} = \sigma_{33}$ and $\sigma_{12} = \sigma_{13}$, Assumption 6 is satisfied.

Finally, We consider the general case of n hypotheses. Now we must consider the multivariate monotone likelihood ratio property, described as follows. For any given $j = 1, \ldots, i_0 - 1$, let $f(x|T_1, \ldots, T_{j-1})$ and $g(x|T_1, \ldots, T_{j-1})$ denote the probability density functions of T_j and T_{i_0} conditional on $T_1, \ldots, T_{j-1}$, respectively.

Assumption 7 (Multivariate Monotone Likelihood Ratio (MMLR)) *For any given values of $T_1, \ldots, T_{j-1}$, $f(x|T_1, \ldots, T_{j-1})$ and $g(x|T_1, \ldots, T_{j-1})$ have the monotone likelihood ratio (MLR) property in x, i.e., for any $x_2 > x_1$, we have*

$$\frac{f(x_2|T_1, \ldots, T_{j-1})}{g(x_2|T_1, \ldots, T_{j-1})} \geq \frac{f(x_1|T_1, \ldots, T_{j-1})}{g(x_1|T_1, \ldots, T_{j-1})}. \tag{13}$$

Theorem 4 *Under Assumptions 2, 3, 5, and 7, the mdFWER of Procedure 2 is strongly controlled at level α.*

6 Conclusions

In this paper, we consider the problem of simultaneously testing multiple prior-ordered hypotheses accompanied by directional decisions. The conventional fixed sequence procedure augmented with additional directional decisions are proved to control the mdFWER under independence and some dependence, whereas, this procedure is also shown to be far too liberal to control the mdFWER, if no dependence assumptions are imposed on the test statistics. We hope that the approaches and techniques developed in this paper will also shed some light on attacking the notoriously challenging problem of controlling the mdFWER under dependence for p-value ordered stepwise procedures.

7 Proofs

PROOF OF LEMMA 1. Let T and P denote the test statistic and the corresponding p-value for testing H, respectively. When testing H, a type 3 error occurs if H is rejected and $\theta T < 0$. Then, the type 3 error rate is given by $Pr(P \leq \alpha, \theta T < 0)$.

When $\theta > 0$, we have

$$\begin{aligned} & Pr(P \leq \alpha, \theta T < 0) = Pr(2F_0(T) \leq \alpha, T < 0) \\ = \; & Pr\left(T \leq F_0^{-1}\left(\frac{\alpha}{2}\right)\right) = F_\theta\left(F_0^{-1}\left(\frac{\alpha}{2}\right)\right) \\ \leq \; & F_0\left(F_0^{-1}\left(\frac{\alpha}{2}\right)\right) = \frac{\alpha}{2}. \end{aligned}$$

The inequality follows from the assumption that F_θ is stochastically increasing in θ. Similarly, when $\theta < 0$, we can also prove that $Pr(P \leq \alpha, \theta T < 0) \leq \frac{\alpha}{2}$. ■

PROOF OF THEOREM 1(i). Induction will be used to show that Procedure 1 strongly controls the mdFWER at level α. First consider the case of $n = 2$. We show control of the mdFWER of Procedure 1 in all possible combinations of true and false null hypotheses while testing two hypotheses H_1 and H_2.

Case I: H_1 is true. Type 1 or type 3 error occurs only when H_1 is rejected.

$$\text{mdFWER} = Pr(P_1 \leq \alpha) \leq \alpha.$$

Case II: Both H_1 and H_2 are false. We have no type 1 errors but only type 3 errors.

$$\begin{aligned}
\text{mdFWER} &= Pr(\{P_1 \leq \alpha, T_1\theta_1 < 0\} \cup \{P_1 \leq \alpha, P_2 \leq \alpha, T_2\theta_2 < 0\}) \\
&\leq Pr(P_1 \leq \alpha, T_1\theta_1 < 0) + Pr(P_2 \leq \alpha, T_2\theta_2 < 0) \\
&\leq \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha.
\end{aligned}$$

The first inequality follows from Bonferroni inequality and the second follows from Lemma 1.

Case III: H_1 is false and H_2 is true. The mdFWER is bounded above by

$$\begin{aligned}
& Pr(\text{ make type 3 error when testing } H_1) + Pr(\text{ make type 1 error when testing } H_2) \\
\leq\ & Pr(P_1 \leq \alpha, T_1\theta_1 < 0) + Pr(P_2 \leq \alpha/2) \\
\leq\ & \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha.
\end{aligned}$$

The first inequality follows from Bonferroni inequality and the second follows from Lemma 1 and $P_2 \sim U(0,1)$ since H_2 is true.

Now assume the inductive hypothesis that the mdFWER is bounded above by α when testing at most $n-1$ hypotheses by using Procedure 1 at level α. In the following, we prove the mdFWER is also bounded above by α when testing n hypotheses $H_1, \ldots, H_n$. Without loss of generality, assume H_1 is a false null (if H_1 is a true null, the desired result directly follows by using the same argument as in Case I of $n = 2$). Then, the mdFWER is bounded above by

$$
\begin{aligned}
& Pr(\text{ make type 3 error when testing } H_1) \\
& \quad + Pr(\text{ make at least one type 1 or type 3 errors when testing } H_2, \ldots, H_n) \\
\leq \; & \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha.
\end{aligned}
$$

The inequality follows from the induction assumption, noticing that $H_2, \ldots, H_n$ are tested by using Procedure 1 at level $\alpha/2$. Thus, the desired result follows.

(ii). We now prove that the critical constants are unimprovable. For instance, when H_1 is true, it is easy to see that the first critical constant, α, is unimprovable. For each given $k = 2, \ldots, n$, when $\theta_i > 0, i = 1, \ldots, k-1$ and $\theta_k = 0$, that is, $H_i, i = 1, \ldots, k-1$ are false and H_k is true, we present a simple joint distribution of the test statistics $T_1, \ldots, T_k$ to show that the kth critical constant of this procedure is also unimprovable.

Define $Z_k \sim N(0,1)$ and $Z_i = \Phi^{-1}(|2\Phi(Z_{i+1}) - 1|), i = 1, \ldots, k-1$, where $\Phi(\cdot)$ is the cdf of N(0, 1). Let q_i denote Z_i's upper $\alpha/2^i$ quantile. It is easy to check that for each $i = 1, \ldots, k$, $Z_i \sim N(0,1)$. Thus, $-q_i$ is Z_i's lower $\alpha/2^i$ quantile. In addition, by the construction of Z_i's, it is easy to see that the event $Z_i \geq q_i$ is equivalent to the event $Z_{i+1} \notin (-q_{i+1}, q_{i+1})$.

Let $T_i = Z_i + \theta_i, i = 1, \ldots, k$, thus $T_i \sim N(\theta_i, 1)$. Then, as $\theta_i \to 0+$ for $i = 1, \ldots, k-1$, we have

$$
\begin{aligned}
\text{mdFWER} &= \sum_{j=1}^{k-1} Pr(T_1 \geq q_1, \ldots, T_{j-1} \geq q_{j-1}, T_j \leq -q_j) \\
& \qquad + Pr(T_1 \geq q_1, \ldots, T_{k-1} \geq q_{k-1}, T_k \notin (-q_k, q_k)) \\
&= \sum_{j=1}^{k-1} Pr(Z_1 \geq q_1, \ldots, Z_{j-1} \geq q_{j-1}, Z_j \leq -q_j) \\
& \qquad + Pr(Z_1 \geq q_1, \ldots, Z_{k-1} \geq q_{k-1}, Z_k \notin (-q_k, q_k)) \\
&= \sum_{j=1}^{k-1} Pr(Z_j \leq -q_j) + Pr(Z_k \notin (-q_k, q_k)) \\
&= \sum_{j=1}^{k-1} \frac{\alpha}{2^j} + \frac{\alpha}{2^{(k-1)}} = \alpha.
\end{aligned}
$$

Thus, the kth critical constant of Procedure 1 is unimprovable and hence each critical constant of Procedure 1 is unimprovable under arbitrary dependence. ■

PROOF OF LEMMA 2. Note that when $\theta_1 > 0$ and $\theta_2 = 0$, we have

$$\begin{aligned}
&\text{mdFWER} \\
= \; & Pr\left(P_1 \leq \alpha, \theta_1 T_1 < 0\right) + Pr\left(P_1 \leq \alpha, \theta_1 T_1 \geq 0, P_2 \leq \alpha\right) \\
= \; & Pr\left(P_1 \leq \alpha, T_1 < 0\right) + Pr\left(P_1 \leq \alpha, T_1 \geq 0, P_2 \leq \alpha, T_2 > 0\right) \\
& + Pr\left(P_1 \leq \alpha, T_1 \geq 0, P_2 \leq \alpha, T_2 \leq 0\right) \\
= \; & Pr\left(2F_0(T_1) \leq \alpha\right) + Pr\left(2(1 - F_0(T_1)) \leq \alpha, 2(1 - F_0(T_2)) \leq \alpha\right) \\
& + Pr\left(2(1 - F_0(T_1)) \leq \alpha, 2F_0(T_2) \leq \alpha\right) \\
= \; & Pr\left(T_1 \leq c_1\right) + Pr\left(T_1 \geq c_2, T_2 \geq c_2\right) + Pr\left(T_1 \geq c_2, T_2 \leq c_1\right) \\
= \; & F_{\theta_1}(c_1) + 1 - F_{\theta_1}(c_2) - F_0(c_2) + F_{(\theta_1,0)}(c_2, c_2) + F_0(c_1) - F_{(\theta_1,0)}(c_2, c_1) \\
= \; & \alpha + F_{\theta_1}(c_1) - F_{\theta_1}(c_2) + F_{(\theta_1,0)}(c_2, c_2) - F_{(\theta_1,0)}(c_2, c_1). \qquad (14)
\end{aligned}$$

Specifically, under Assumption 1 (independence), (14) can be simplified as,

$$\begin{aligned}
& \alpha + F_{\theta_1}(c_1) - F_{\theta_1}(c_2) + F_{\theta_1}(c_2)F_0(c_2) - F_{\theta_1}(c_2)F_0(c_1) \\
= \; & \alpha + F_{\theta_1}(c_1) - \alpha F_{\theta_1}(c_2).
\end{aligned}$$

Similarly, when $\theta_1 < 0$ and $\theta_2 = 0$, we can prove that

$$\text{mdFWER} = 1 + F_{\theta_1}(c_1) - F_{\theta_1}(c_2) + F_{(\theta_1,0)}(c_1, c_1) - F_{(\theta_1,0)}(c_1, c_2). \blacksquare$$

PROOF OF LEMMA 3. By using the same arguments as in Theorem 1, we can easily prove control of the mdFWER of Procedure 2 in the case of $n = 2$ when H_1 is true or both H_1 and H_2 are false. In the following, we prove the desired result also holds when H_1 is false and H_2 is true.

Note that H_1 is false and H_2 is true imply $\theta_1 \neq 0$ and $\theta_2 = 0$. To show that the mdFWER is controlled for $\theta_1 > 0$ and $\theta_2 = 0$, we only need to show by Lemma 2 that

$\alpha + F_{\theta_1}(c_1) - \alpha F_{\theta_1}(c_2) \leq \alpha$. This is equivalent to show

$$F_{\theta_1}(c_2)\left(F_0(c_2) - F_0(c_1)\right) \quad \leq \quad F_{\theta_1}(c_2) - F_{\theta_1}(c_1). \tag{15}$$

For proving (15), it is enough to prove the following, as $0 \leq F_0(c_2) \leq 1$,

$$F_{\theta_1}(c_2)\left(F_0(c_2) - F_0(c_1)\right) \leq F_0(c_2)\left(F_{\theta_1}(c_2) - F_{\theta_1}(c_1)\right). \tag{16}$$

Dividing both sides of (16) by $F_{\theta_1}(c_2)F_0(c_2)$, we see that we only need to prove,

$$1 - \frac{F_0(c_1)}{F_0(c_2)} \quad \leq \quad 1 - \frac{F_{\theta_1}(c_1)}{F_{\theta_1}(c_2)},$$

which follows directly from (5) and Assumption 2 (MLR).

Similarly, to show that the mdFWER is controlled for $\theta_1 < 0$ and $\theta_2 = 0$, we only need to show by Lemma 2 that $1 + \alpha F_{\theta_1}(c_1) - F_{\theta_1}(c_2) \leq \alpha$. This is equivalent to showing

$$(1-\alpha)\left(1 - F_{\theta_1}(c_1)\right) \leq F_{\theta_1}(c_2) - F_{\theta_1}(c_1).$$

Writing $1-\alpha$ as $(1 - F_0(c_1)) - (1 - F_0(c_2))$ and writing $F_{\theta_1}(c_2) - F_{\theta_1}(c_1)$ as $(1 - F_{\theta_1}(c_1)) - (1 - F_{\theta_1}(c_2))$, we get that it is equivalent to prove

$$\left[(1 - F_0(c_1)) - (1 - F_0(c_2))\right]\left(1 - F_{\theta_1}(c_1)\right) \leq \left(1 - F_{\theta_1}(c_1)\right) - \left(1 - F_{\theta_1}(c_2)\right). \tag{17}$$

Since $0 \leq 1 - F_0(c_1) \leq 1$, to prove inequality (17), it is enough to prove the following,

$$\begin{aligned} & \left(1 - F_{\theta_1}(c_1)\right)\left[(1 - F_0(c_1)) - (1 - F_0(c_2))\right] \\ \leq \quad & \left(1 - F_0(c_1)\right)\left[1 - F_{\theta_1}(c_1)\right] - \left[1 - F_{\theta_1}(c_2)\right]. \end{aligned} \tag{18}$$

Dividing both sides of (18) by $\left(1 - F_{\theta_1}(c_1)\right)\left(1 - F_0(c_1)\right)$, we see that proving (17) is equivalent to showing

$$\frac{1 - F_{\theta_1}(c_2)}{1 - F_{\theta_1}(c_1)} \leq \frac{1 - F_0(c_2)}{1 - F_0(c_1)}, \tag{19}$$

which follows directly from (6) and Assumption 2 (MLR).

By combining the discussion of the above two cases, the desired result follows. ∎

PROOF OF THEOREM 2. The proof is by induction on number of hypotheses n. We already proved strong control of the mdFWER for $n = 2$ in Lemma 3. Let us assume the result holds for testing any $n = k$ hypotheses, that is, mdFWER $\leq \alpha$ while testing any k pre-ordered hypotheses. We now argue that is will hold for $n = k+1$ hypotheses. Without loss of generality, assume H_1 is a false null, as in the proof of Theorem 1.

Let $V_{k+1}^{(-1)}$ denote the total number of type 1 or type 3 errors committed while testing $H_2, \ldots, H_{k+1}$ and excluding H_1. Then, by the inductive hypothesis, the mdFWER while testing the k hypotheses $H_2, \ldots, H_{k+1}$ is $Pr(V_{k+1}^{(-1)} > 0) \leq \alpha$. Then, the mdFWER of testing $k+1$ hypotheses $H_1, \ldots, H_{k+1}$ is defined by

$$
\begin{aligned}
& Pr\left(\{P_1 \leq \alpha, T_1\theta_1 < 0\} \cup \{P_1 \leq \alpha, T_1\theta_1 \geq 0, V_{k+1}^{(-1)} > 0\}\right) \\
= \;& Pr\left(P_1 \leq \alpha, T_1\theta_1 < 0\right) + Pr\left(P_1 \leq \alpha, T_1\theta_1 \geq 0\right) \cdot Pr\left(V_{k+1}^{(-1)} > 0\right) \\
\leq \;& Pr\left(P_1 \leq \alpha, T_1\theta_1 < 0\right) + \alpha \, Pr\left(P_1 \leq \alpha, T_1\theta_1 \geq 0\right). \qquad (20)
\end{aligned}
$$

The equality follows by Assumption 1 (independence) and the inequality follows by the inductive hypothesis. Note that (20) is the same as (8) under independence, which is equal to the mdFWER of Procedure 2 in the case of two hypotheses. So again by applying Lemma 3, we get that mdFWER $\leq \alpha$ for $n = k+1$. Hence, the proof follows by induction. ∎

PROOF OF THEOREM 3 . Without loss of generality, we assume $\theta_i > 0$ if $\theta_i \neq 0$ for $i = 1, \ldots, n$. Also, if there exists an i with $\theta_i = 0$, by induction, we can simply assume $i_0 = n$. Thus, to prove the mdFWER control of Procedure 2, we only need to consider two cases:

(i) $\theta_i > 0$ for $i = 1, \ldots, n$;

(ii) $\theta_i > 0$ for $i = 1, \ldots, n-1$ and $\theta_n = 0$.

Case (i). Consider the general case of $\theta_i > 0, i = 1, \ldots, n$. By Assumption 3, the test statistics $T_1, \ldots, T_n$ are positively regression dependent. For $j = 1, \ldots, n-1$, let E_{n-j} denote the event of making at least one type 3 error when testing $H_{j+1}, \ldots, H_n$ using

Procedure 2 at level α. By using induction, we prove the following two lemmas hold.

Lemma 4 *Assume the conditions of Theorem 3. For $j = 1, \ldots, n-1$, the following inequality holds.*

$$Pr(E_{n-j}|T_1 > c_2, \ldots, T_j > c_2) \leq \alpha. \tag{21}$$

PROOF OF LEMMA 4. We prove the result by using reverse induction. When $j = n-1$, we have

$$\begin{aligned}
& Pr(E_{n-j}|T_1 > c_2, \ldots, T_j > c_2) \\
= \; & Pr(T_n < c_1|T_1 > c_2, \ldots, T_{n-1} > c_2) \\
= \; & \frac{Pr(T_n < c_1)Pr(T_1 > c_2, \ldots, T_{n-1} > c_2|T_n < c_1)}{Pr(T_1 > c_2, \ldots, T_{n-1} > c_2)} \\
\leq \; & Pr(T_n < c_1) \leq \alpha.
\end{aligned}$$

The inequality follows from Assumption 3.

Assume the inequality (21) holds for $j = m$. In the following, we prove that it also holds for $j = m-1$. Note that

$$\begin{aligned}
& Pr(E_{n-m+1}|T_1 > c_2, \ldots, T_{m-1} > c_2) \\
= \; & Pr\left(\{T_m < c_1\} \bigcup \left(\{T_m > c_2\} \bigcap E_{n-m}\right) \Big| T_1 > c_2, \ldots, T_{m-1} > c_2\right) \\
= \; & Pr\left(T_m < c_1 \big| T_1 > c_2, \ldots, T_{m-1} > c_2\right) + Pr\left(\{T_m > c_2\} \bigcap E_{n-m} \big| T_1 > c_2, \ldots, T_{m-1} > c_2\right) \\
= \; & Pr\left(T_m < c_1 \big| T_1 > c_2, \ldots, T_{m-1} > c_2\right) \\
& + Pr\left(T_m > c_2 \big| T_1 > c_2, \ldots, T_{m-1} > c_2\right) Pr\left(E_{n-m} \big| T_1 > c_2, \ldots, T_m > c_2\right) \\
\leq \; & Pr\left(T_m < c_1 \big| T_1 > c_2, \ldots, T_{m-1} > c_2\right) + \alpha Pr\left(T_m > c_2 \big| T_1 > c_2, \ldots, T_{m-1} > c_2\right) \\
\leq \; & \alpha.
\end{aligned}$$

Therefore, the desired result follows. Here, the first inequality follows from the assumption of induction and the second follows from Lemma 5 below. ■

Lemma 5 *Assume the conditions of Theorem 3. For $j = 1, \ldots, n-1$, the following* *inequality holds:*

$$Pr\left(T_j < c_1 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right) + \alpha Pr\left(T_j > c_2 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right) \leq \alpha. \tag{22}$$

Specifically, for $j = 1$, we have

$$Pr\left(T_1 < c_1\right) + \alpha Pr\left(T_1 > c_2\right) \leq \alpha.$$

PROOF OF LEMMA 5. To prove the inequality (22), it is enough to show that

$$Pr\left(T_j < c_1 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right) \leq \alpha Pr\left(T_j < c_2 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right),$$

which is equivalent to

$$\begin{aligned} & (1-\alpha) Pr\left(T_j < c_2 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right) \\ \leq\ & Pr\left(T_j < c_2 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right) - Pr\left(T_j < c_1 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right). \end{aligned}$$

Note that

$$1 - \alpha = Pr_{\theta_j = 0}(T_j < c_2) - Pr_{\theta_j = 0}(T_j < c_1).$$

Thus, the above inequality is equivalent to

$$Pr_{\theta_j=0}(T_j < c_2) - Pr_{\theta_j=0}(T_j < c_1) \leq 1 - \frac{Pr\left(T_j < c_1 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right)}{Pr\left(T_j < c_2 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right)},$$

which in turn is implied by

$$1 - \frac{Pr_{\theta_j=0}(T_j < c_1)}{Pr_{\theta_j=0}(T_j < c_2)} \leq 1 - \frac{Pr\left(T_j < c_1 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right)}{Pr\left(T_j < c_2 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right)}. \tag{23}$$

Note that by Assumption 2, we have

$$\frac{Pr(T_j < c_1)}{Pr(T_j < c_2)} \leq \frac{Pr_{\theta_j=0}(T_j < c_1)}{Pr_{\theta_j=0}(T_j < c_2)}.$$

Thus, to prove the inequality (23), we only need to show that

$$\frac{Pr\left(T_j < c_1 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right)}{Pr\left(T_j < c_2 \middle| T_1 > c_2, \ldots, T_{j-1} > c_2\right)} \leq \frac{Pr(T_j < c_1)}{Pr(T_j < c_2)},$$

which is equivalent to

$$Pr\left(T_1 > c_2, \ldots, T_{j-1} > c_2 \middle| T_j < c_1\right) \leq Pr\left(T_1 > c_2, \ldots, T_{j-1} > c_2 \middle| T_j < c_2\right),$$

which follows from Assumption 3. Therefore, the desired result follows. ■

Based on Lemmas 4 and 5, we have

$$\begin{aligned}
& \text{mdFWER} = Pr(T_1 < c_1) + \sum_{j=2}^{n} Pr(T_1 > c_2, \ldots, T_{j-1} > c_2, T_j < c_1) \\
= \;& Pr(T_1 < c_1) + Pr(T_1 > c_2) \sum_{j=2}^{n} Pr(T_2 > c_2, \ldots, T_{j-1} > c_2, T_j < c_1 | T_1 > c_2) \\
= \;& Pr(T_1 < c_1) + Pr(T_1 > c_2) Pr(E_{n-1} | T_1 > c_2) \\
\leq \;& Pr(T_1 < c_1) + \alpha Pr(T_1 > c_2) \\
\leq \;& \alpha.
\end{aligned}$$

Therefore, the mdFWER is controlled at level α for Case (i). Here, the first inequality follows from Lemma 4 and the second follows from Lemma 5.

Case (ii). Consider the general case of $\theta_i > 0, i = 1, \ldots, n-1$ and $\theta_n = 0$. Under Assumption 3, $T_i, i = 1, \ldots, n-1$ are positively regression dependent and under Assumption 4, T_n is independent of T_i's . Note that

$$\begin{aligned}
& \text{mdFWER} \\
= \;& \sum_{j=1}^{n-1} Pr(T_1 > c_2, \ldots, T_{j-1} > c_2, T_j < c_1) \\
& \quad + Pr(T_1 > c_2, \ldots, T_{n-1} > c_2, T_n < c_1) + \; Pr(T_1 > c_2, \ldots, T_n > c_2) \\
= \;& \sum_{j=1}^{n-1} Pr(T_1 > c_2, \ldots, T_{j-1} > c_2, T_j < c_1) + \alpha Pr(T_1 > c_2, \ldots, T_{n-1} > c_2).
\end{aligned}$$

The second equality follows from Assumption 4.

For $m = 1, \ldots, n-1$, define

$$\Delta_m = \sum_{j=1}^{m} Pr(T_1 > c_2, \ldots, T_{j-1} > c_2, T_j < c_1) + \alpha Pr(T_1 > c_2, \ldots, T_m > c_2).$$

Thus, mdFWER $= \Delta_{n-1}$. By using induction, we prove below that $\Delta_m \leq \alpha$ for $m = 1, \ldots, n-1$.

For $m = 1$, by using Lemma 5, we have

$$\Delta_1 = Pr\left(T_1 < c_1\right) + \alpha Pr\left(T_1 > c_2\right) \leq \alpha.$$

Assume $\Delta_m \leq \alpha$. In the following, we show $\Delta_{m+1} \leq \alpha$. Note that

$$\begin{aligned}
\Delta_{m+1} &= \sum_{j=1}^{m+1} Pr(T_1 > c_2, \ldots, T_{j-1} > c_2, T_j < c_1) \\
&\quad + \alpha Pr(T_1 > c_2, \ldots, T_m > c_2, T_{m+1} > c_2) \\
&= \sum_{j=1}^{m} Pr(T_1 > c_2, \ldots, T_{j-1} > c_2, T_j < c_1) \\
&\quad + Pr(T_1 > c_2, \ldots, T_m > c_2)\left[Pr(T_{m+1} < c_1 | T_1 > c_2, \ldots, T_m > c_2)\right. \\
&\qquad \left. + \alpha Pr(T_{m+1} > c_2 | T_1 > c_2, \ldots, T_m > c_2)\right] \\
&\leq \sum_{j=1}^{m} Pr(T_1 > c_2, \ldots, T_{j-1} > c_2, T_j < c_1) + \alpha Pr(T_1 > c_2, \ldots, T_m > c_2) \\
&= \Delta_m \leq \alpha. \qquad (24)
\end{aligned}$$

The first inequality follows from Lemma 5 and the second follows from the inductive hypothesis. Thus, $\Delta_m \leq \alpha$ for $m = 1, \ldots, n-1$. Therefore, mdFWER $= \Delta_{n-1} \leq \alpha$, the desired result.

Combining the arguments of Cases (i) and (ii), the proof of Theorem 3 is complete. ■

PROOF OF PROPOSITION 1. From the proof of Theorem 1 and by Lemma 1, it is easy to see that we only need to prove the mdFWER control of Procedure 2 when H_1 is false and H_2 is true, i.e., $\theta_1 \neq 0$ and $\theta_2 = 0$.

Case I: $\theta_1 > 0$ **and** $\theta_2 = 0$**.** By Lemma 2, the mdFWER of Procedure 2 is controlled at level α if we have the following:

$$F_{\theta_1}(c_1) - F_{\theta_1}(c_2) + F_{(\theta_1,0)}(c_2, c_2) - F_{(\theta_1,0)}(c_2, c_1) \leq 0.$$

After rewriting $F_{(\theta_1,0)}(x, y)$ as $Pr(T_1 \leq x, T_2 \leq y)$ and then dividing through by $Pr(T_1 \leq c_2)$, we get,

$$Pr\left(T_2 \leq c_2 | T_1 \leq c_2\right) - Pr\left(T_2 \leq c_1 | T_1 \leq c_2\right) \leq 1 - \frac{Pr(T_1 \leq c_1)}{Pr(T_1 \leq c_2)}.$$

Dividing by $Pr\left(T_2 \leq c_2 | T_1 \leq c_2\right)$, we get,

$$1 - \frac{Pr\left(T_2 \leq c_1 | T_1 \leq c_2\right)}{Pr\left(T_2 \leq c_2 | T_1 \leq c_2\right)} \leq \frac{1}{Pr\left(T_2 \leq c_2 | T_1 \leq c_2\right)} \left(1 - \frac{Pr(T_1 \leq c_1)}{Pr(T_1 \leq c_2)}\right). \tag{25}$$

For proving (25), it is enough to prove the following inequality, as $\frac{1}{Pr(T_2 \leq c_2 | T_1 \leq c_2)} \geq 1$.

$$1 - \frac{Pr\left(T_2 \leq c_1 | T_1 \leq c_2\right)}{Pr\left(T_2 \leq c_2 | T_1 \leq c_2\right)} \leq 1 - \frac{Pr(T_1 \leq c_1)}{Pr(T_1 \leq c_2)}. \tag{26}$$

By Assumption 2 and (5), it follows that $\frac{F_0(c_2)}{F_0(c_1)} \leq \frac{F_{\theta_1}(c_2)}{F_{\theta_1}(c_1)}$, which is equivalent to, $1 - \frac{Pr(T_2 \leq c_1)}{Pr(T_2 \leq c_2)} \leq 1 - \frac{Pr(T_1 \leq c_1)}{Pr(T_1 \leq c_2)}$. Thus for proving (25), it is enough to prove the following:

$$1 - \frac{Pr\left(T_2 \leq c_1 | T_1 \leq c_2\right)}{Pr\left(T_2 \leq c_2 | T_1 \leq c_2\right)} \leq 1 - \frac{Pr(T_2 \leq c_1)}{Pr(T_2 \leq c_2)}\,. \tag{27}$$

But, (27) is equivalent to showing

$$Pr\left(T_1 \leq c_2 | T_2 \leq c_1\right) \geq Pr\left(T_1 \leq c_2 | T_2 \leq c_2\right),$$

which follows directly from Assumption 5.

Case II: $\theta_1 < 0$ **and** $\theta_2 = 0$**.** Similarly, by Lemma 2, the mdFWER of Procedure 2 is controlled at level α if we have the following:

$$1 + F_{\theta_1}(c_1) - F_{\theta_1}(c_2) + F_{(\theta_1,0)}(c_1, c_1) - F_{(\theta_1,0)}(c_1, c_2) \leq \alpha, \tag{28}$$

which after some rearrangement and rewriting $1-\alpha$ as $F_0(c_2)-F_0(c_1)$ gives,

$$\left(F_0(c_2)-F_{(\theta_1,0)}(c_1,c_2)\right)-\left(F_0(c_1)-F_{(\theta_1,0)}(c_1,c_1)\right)\leq(1-F_{\theta_1}(c_1))-(1-F_{\theta_1}(c_2))\,.\text{(29)}$$

Thus, proving (28) is equivalent to proving that

$$Pr\left(T_1\geq c_1,T_2\leq c_2\right)-Pr\left(T_1\geq c_1,T_2\leq c_1\right)\leq Pr\left(T_1\geq c_1\right)-Pr\left(T_1\geq c_2\right).$$

Dividing through by $Pr(T_1\geq c_1)$, we get

$$Pr\left(T_2\geq c_1|T_1\geq c_1\right)-Pr\left(T_2\geq c_2|T_1\geq c_1\right)\leq 1-\frac{Pr(T_1\geq c_2)}{Pr(T_1\geq c_1)}. \tag{30}$$

Thus to prove (28), it is enough to prove the following,

$$1-\frac{Pr\left(T_2\geq c_2|T_1\geq c_1\right)}{Pr\left(T_2\geq c_1|T_1\geq c_1\right)}\leq 1-\frac{Pr(T_1\geq c_2)}{Pr(T_1\geq c_1)},$$

which is equivalent to proving,

$$\frac{Pr\left(T_2\geq c_2|T_1\geq c_1\right)}{Pr\left(T_2\geq c_1|T_1\geq c_1\right)}\geq\frac{Pr(T_1\geq c_2)}{Pr(T_1\geq c_1)}. \tag{31}$$

By Assumption 2 and (6), it follows that for $\theta_1<0$, $\frac{Pr(T_1\geq c_2)}{Pr(T_1\geq c_1)}\leq\frac{Pr(T_2\geq c_2)}{Pr(T_2\geq c_1)}$. Thus to prove (28), it is enough to prove the following,

$$\frac{Pr\left(T_2\geq c_2|T_1\geq c_1\right)}{Pr\left(T_2\geq c_1|T_1\geq c_1\right)}\geq\frac{Pr(T_2\geq c_2)}{Pr(T_2\geq c_1)}. \tag{32}$$

But (32) is equivalent to showing

$$Pr\left(T_1\geq c_1|T_2\geq c_2\right)\geq Pr\left(T_1\geq c_1|T_2\geq c_1\right), \tag{33}$$

which follows directly from Assumption 5. By combining the arguments of the above two cases, the desired result follows. ■

PROOF OF PROPOSITION 2. By Corollary 1, without loss of generality, assume that $\theta_i >$

$0, i = 1, 2$ and $\theta_3 = 0$, that is, H_1 and H_2 are false and H_3 is true. Note that

$$\begin{aligned}
& \text{mdFWER} && (34)\\
= \;& Pr(T_1 \leq c_1) + Pr(T_1 \geq c_2, T_2 \leq c_1) + Pr\left(T_1 \geq c_2, T_2 \geq c_2, T_3 \notin (c_1, c_2)\right).
\end{aligned}$$

In the following, we prove that

$$\begin{aligned}
& Pr(T_1 \geq c_2, T_2 \leq c_1) + Pr\left(T_1 \geq c_2, T_2 \geq c_2, T_3 \notin (c_1, c_2)\right)\\
\leq \;& Pr\left(T_1 \geq c_2, T_3 \notin (c_1, c_2)\right). && (35)
\end{aligned}$$

To prove (35), it is enough to show the following inequality:

$$Pr(T_2 \leq c_1 | T_1) + Pr\left(T_2 \geq c_2, T_3 \notin (c_1, c_2) | T_1\right) \leq Pr\left(T_3 \notin (c_1, c_2) | T_1\right). \quad (36)$$

Note that

$$Pr\left(T_2 \geq c_2, T_3 \leq c_1 | T_1\right) = Pr(T_3 \leq c_1 | T_1) - Pr\left(T_2 < c_2, T_3 \leq c_1 | T_1\right) \quad (37)$$

and

$$\begin{aligned}
& Pr\left(T_2 \geq c_2, T_3 \geq c_2 | T_1\right)\\
= \;& 1 - Pr(T_2 < c_2 | T_1) - Pr(T_3 < c_2 | T_1) + Pr\left(T_2 < c_2, T_3 < c_2 | T_1\right). && (38)
\end{aligned}$$

In addition, we have

$$Pr\left(T_3 \notin (c_1, c_2) | T_1\right) = 1 + Pr(T_3 \leq c_1 | T_1) - Pr(T_3 < c_2 | T_1). \quad (39)$$

Thus, in order to show (36), by combining (37)-(39), we only need to prove the following inequality:

$$\begin{aligned}
& Pr\left(T_2 < c_2, T_3 < c_2 | T_1\right) - Pr\left(T_2 < c_2, T_3 \leq c_1 | T_1\right)\\
\leq \;& Pr(T_2 < c_2 | T_1) - Pr(T_2 \leq c_1 | T_1). && (40)
\end{aligned}$$

Note that (40) can be rewritten as

$$
\begin{aligned}
& Pr\left(T_2 < c_2, T_3 < c_2|T_1\right)\left[1 - \frac{Pr\left(T_2 < c_2, T_3 \le c_1|T_1\right)}{Pr\left(T_2 < c_2, T_3 < c_2|T_1\right)}\right] \\
\le \; & Pr(T_2 < c_2|T_1)\left[1 - \frac{Pr(T_2 \le c_1|T_1)}{Pr(T_2 < c_2|T_1)}\right]. \qquad (41)
\end{aligned}
$$

Thus, to prove (40), it is enough to show

$$1 - \frac{Pr\left(T_2 < c_2, T_3 \le c_1|T_1\right)}{Pr\left(T_2 < c_2, T_3 < c_2|T_1\right)} \le 1 - \frac{Pr(T_2 \le c_1|T_1)}{Pr(T_2 < c_2|T_1)}. \qquad (42)$$

That is,

$$\frac{Pr(T_2 \le c_1|T_1)}{Pr(T_2 < c_2|T_1)} \le \frac{Pr\left(T_2 < c_2, T_3 \le c_1|T_1\right)}{Pr\left(T_2 < c_2, T_3 < c_2|T_1\right)}. \qquad (43)$$

By Assumption 6 (BMLR), we have

$$\frac{Pr(T_2 \le x_2|T_1)}{Pr(T_3 \le x_2|T_1)} \ge \frac{Pr(T_2 \le x_1|T_1)}{Pr(T_3 \le x_1|T_1)}. \qquad (44)$$

By (44), to prove (43), it is enough to show

$$\frac{Pr(T_3 \le c_1|T_1)}{Pr(T_3 < c_2|T_1)} \le \frac{Pr\left(T_2 < c_2, T_3 \le c_1|T_1\right)}{Pr\left(T_2 < c_2, T_3 < c_2|T_1\right)}. \qquad (45)$$

That is,

$$Pr\left(T_2 < c_2|T_3 < c_2, T_1\right) \le Pr\left(T_2 < c_2|T_3 < c_1, T_1\right). \qquad (46)$$

The inequality (46) holds under Assumption 5. Therefore, the inequality (35) holds.

Based on (34)-(35) and Proposition 1, we have

$$\text{mdFWER} = Pr(T_1 \le c_1) + Pr\left(T_1 \ge c_2, T_3 \notin (c_1, c_2)\right) \le \alpha.$$

Thus, the desired result follows. ■

PROOF OF THEOREM 4. By Corollary 1, without loss of generality, assume that $\theta_i > 0, i =$

$1, \ldots, n-1$ and $\theta_n = 0$, that is, $H_i, i = 1, \ldots, n-1$ are false and H_n is true. Note that

$$\begin{aligned} & \text{mdFWER} \quad (47) \\ = & \sum_{j=1}^{n-1} Pr(T_1 \geq c_2, \ldots, T_{j-1} \geq c_2, T_j \leq c_1) + Pr(T_1 \geq c_2, \ldots, T_{n-1} \geq c_2, T_n \notin (c_1, c_2)). \end{aligned}$$

In the following, we prove that

$$\begin{aligned} & Pr(T_1 \geq c_2, \ldots, T_{n-2} \geq c_2, T_{n-1} \leq c_1) + Pr\left(T_1 \geq c_2, \ldots, T_{n-1} \geq c_2, T_n \notin (c_1, c_2)\right) \\ \leq & \; Pr\left(T_1 \geq c_2, \ldots, T_{n-2} \geq c_2, T_n \notin (c_1, c_2)\right). \end{aligned} \quad (48)$$

To prove (48), it is enough to show the following inequality:

$$\begin{aligned} & Pr(T_{n-1} \leq c_1 | T_1, \ldots, T_{n-2}) + Pr\left(T_{n-1} \geq c_2, T_n \notin (c_1, c_2) | T_1, \ldots, T_{n-2}\right) \\ \leq & \; Pr\left(T_n \notin (c_1, c_2) | T_1, \ldots, T_{n-2}\right). \end{aligned} \quad (49)$$

By using the same argument as in proving (36) in the case of three hypotheses, we can prove that the inequality (49) holds under Assumptions 5 and 7. Then, by combining (47) and (48), we have

$$\begin{aligned} & \text{mdFWER} \quad (50) \\ \leq & \sum_{j=1}^{n-2} Pr(T_1 \geq c_2, \ldots, T_{j-1} \geq c_2, T_j \leq c_1) + Pr(T_1 \geq c_2, \ldots, T_{n-2} \geq c_2, T_n \notin (c_1, c_2)). \end{aligned}$$

Note that the right-hand side of (50) is the mdFWER of Procedure 2 when testing $H_1, \ldots, H_{n-2}$, H_n. By induction and Proposition 1, the mdFWER is bounded above by α, the desired result.

Acknowledgements

The research of Wenge Guo was supported in part by NSF Grant DMS-1309162 and the research of Joseph Romano was supported in part by NSF Grant DMS-0707085.

References

[1] BENJAMINI, Y. and YEKUTIELI, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.* **29**, 1165-1188.

[2] DMITRIENKO, A., TAMHANE, A. and BRETZ, F. (2009). *Multiple Testing Problems in Pharmaceutical Statistics.* Chapman and Hall/CRC Press, New York.

[3] DMITRIENKO, A., D'AGOSTINO, R. and HUQUE, M. (2013). Key multiplicity issues in clinical drug development. *Statistics in Medicine* **32**, 1079–1111.

[4] FINNER, H. (1994). Testing multiple hypotheses: general theory, specific problems, and relationships to other multiple decision procedures. Habilitationsschrift, Fachbereich IV Mathematik, Univ. Trier.

[5] FINNER, H. (1999). Stepwise multiple test procedures and control of directional errors. *Ann. Statist.* **27**, 274-289.

[6] GUO, W. and ROMANO, J. (2015). On stepwise control of directional errors under independence and some dependence. *Journal of Statistical Planning and Inference*, in press.

[7] HOCHBERG Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* **75**, 800-802.

[8] HOLM, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* **6**, 65-70.

[9] LIU, W. (1997). Control of directional errors with step-up multiple tests. *Statist. Probab. Lett.* **31**, 239-242.

[10] MAURER W., HOTHORN L. and LEHMACHER W. (1995). Multiple comparisons in drug clinical trials and preclinical assays: a-priori ordered hypotheses. In *Biometrie in der Chemisch-pharmazeutischen Industrie,* J. Vollmar, eds. **6**:3–18, Fischer Verlag, Stuttgart.

[11] SARKAR, S. (2002). Some results on false discovery rate in stepwise multiple testing procedures. *Ann. Statist.* **30** 239-257.

[12] SARKAR, S., SEN, P. K. and FINNER, H. (2004). On two results in multiple testing. In *Recent Developments in Multiple Comparisons.* IMS Lectures Notes-Monograph Series, 47, Y. Benjamini, F. Bretz and S. Sarkar, eds. 89-99, Institute of Mathematical Statistics, Beachwood.

[13] SARKAR, S. and GUO, W. (2010). Procedures controlling generalized false discovery rate using bivariate distributions of the null p-values. *Statistica Sinica*, **20**, 1227-1238.

[14] SHAFFER, J. P. (1980). Control of directional errors with stagewise multiple test procedures. *Ann. Statist.* **8**, 1342-1347.

[15] SHAFFER, J. P. (2002). Multiplicity, directional (type III) errors, and the null hypothesis. *Psychological Methods* **7**, 356-369.

[16] WIENS B. (2003). A fixed-sequence Bonferroni procedure for testing multiple endpoints. *Pharmaceutical Statistics* **2**, 211-215.

[17] WIENS B. and DMITRIENKO A. (2005). The fallback procedure for evaluating a single family of hypotheses. *Journal of Biopharmaceutical Statistics* **15**, 929-942.