Web-Based Supplementary Materials for "Multiple McNemar Tests,"

by **Peter H. Westfall,**[1,*], **James F. Troendle,**[2] **and Gene Pennello**[3]

[1]Area of ISQS, Texas Tech University, Lubbock, TX 79409-2101, USA

[2]Biostatistics and Bioinformatics Branch, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, Bethesda, MD 20892, USA

[3]US Food and Drug Administration, Epidemiology Branch, Center for Devices and Radiological Health, HFZ-541, 1350 Piccard Drive, 20850 Rockville, MD, USA

[*]*email:* peter.westfall@ttu.edu

This document provides proofs and additional simulation results to supplement the above-titled paper by the above indicated authors.

1 Theorems and Proofs

Theorem 1 *The test that rejects H_I when $T_I(\boldsymbol{N}_I^+) \leq c_I^\alpha(\boldsymbol{N}_I)$ has type I error rate $\leq \alpha$.*

Proof. First, define $q_\ell(\alpha, \boldsymbol{n}_I; H_\ell) = P_{H_\ell}\left\{p_\ell(N_{01}^{(\ell)}, N_d^{(\ell)}) \leq c_I^\alpha(\boldsymbol{n}_I) \mid N_d^{(\ell)} = n_d^{(\ell)}\right\}$. Note that $N_{01}^{(\ell)}|N_d^{(\ell)} = n_d^{(\ell)} \sim B(n_d^{(\ell)}, 0.5)$ whenever H_ℓ is true; in particular, if $\ell \in I$, then $N_{01}^{(\ell)}|N_d^{(\ell)} = n_d^{(\ell)} \sim B(n_d^{(\ell)}, 0.5)$ when H_I is true. Hence when $\ell \in I$,

$$q_\ell(\alpha, \boldsymbol{n}_I; H_I) \equiv P_{H_I}\left\{p_\ell(N_{01}^{(\ell)}, N_d^{(\ell)}) \leq c_I^\alpha(\boldsymbol{n}_I) \mid N_d^{(\ell)} = n_d^{(\ell)}\right\} = q_\ell(\alpha, \boldsymbol{n}_I; H_\ell). \qquad (1^*)$$

Now, by (7, main article), $\sum_{\ell \in I} q_\ell(\alpha, \boldsymbol{n}_I; H_\ell) \leq \alpha$ for all observable $\boldsymbol{n}_I$; implying $\sum_{\ell \in I} q_\ell(\alpha, \boldsymbol{N}_I; H_\ell) \leq$

α with probability 1.0. By (8, main article), $\sum_{\ell \in I} q_\ell(\alpha, \boldsymbol{N}_I; H_I) \leq \alpha$ with probability 1.0 as well.

Hence

$$\begin{aligned}
P_{H_I}\{T_I(\boldsymbol{N}_I^+) \leq c_I^\alpha(\boldsymbol{N}_I)\} & \\
&= P_{H_I}\left\{\min_{\ell \in I} p_\ell(N_{01}^{(\ell)}, N_d^{(\ell)}) \leq c_I^\alpha(\boldsymbol{N}_I)\right\} \\
&\leq \sum_{\ell \in I} P_{H_I}\left\{p_\ell(N_{01}^{(\ell)}, N_d^{(\ell)}) \leq c_I^\alpha(\boldsymbol{N}_I)\right\} \qquad \text{by Boole's inequality} \\
&= \sum_{\ell \in I} E_{H_I}\left\{q_\ell(\alpha, \boldsymbol{N}_I; H_I)\right\} \qquad \text{by (1*)} \\
&= E_{H_I} \sum_{\ell \in I} q_\ell(\alpha, \boldsymbol{N}_I; H_I) \\
&\leq \alpha
\end{aligned}$$

■

Theorem 2 *If the assumption and null hypotheses are as given in Section 3 of the main article, then the discrete Bonferroni method controls the FWER in the strong sense.*

Proof. Suppose the set of true nulls is I; then $FWER = P(\text{Reject } H_\ell \text{ for some } \ell \in I\ )$. Now, $\{\text{Reject } H_\ell \text{ for some } \ell \in I\} \subseteq \{T_I(\boldsymbol{N}_I^+) \leq c_{I'}^\alpha(\boldsymbol{N}_{I'}) \text{ for some } I' \supseteq I\}$. Note further that the critical values $c_I^\alpha(\boldsymbol{N}_I)$ are monotonic in subset size: if $I' \supseteq I$ then $c_{I'}^\alpha(\boldsymbol{N}_{I'}) \leq c_I^\alpha(\boldsymbol{N}_I)$. Hence

$$\begin{aligned}
P_I(\text{Reject } H_\ell \text{ for some } \ell \in I) &\leq P_I\{T_I(\boldsymbol{N}_I^+) \leq c_{I'}^\alpha(\boldsymbol{N}_{I'}) \text{ for some } I' \supseteq I\} \\
&\leq P_I\{T_I(\boldsymbol{N}_I^+) \leq c_I^\alpha(\boldsymbol{N}_I)\} \quad \text{(by monotonicity)} \\
&\leq \alpha \quad \text{(by Theorem 1)}
\end{aligned}$$

■

2 Additional Simulations

The following tables provide additional simulation results to supplement the paper. For details of the simulation study, refer to Section 9 of the main article.

Web Table1A: Power comparison of the procedures, all pairwise comparisons case, $n = 25$.

			Power			
μ'	Covariance	$E(\rho^2)$	Disc, Ex	Holm, Ex	Holm, McN	Boot, (5)
$[0^{\times 4}, 1]$	any	0	.36	.26	.34	.42
$[0^{\times 4}, 1]$	CS	.5	.67	.44	.58	.66
$[0^{\times 4}, 1]$	FA^{+}	.5	.66	.42	.56	.66
$[0^{\times 4}, 1]$	$FA^{+/-}$	.5	.74	.47	.62	.76
$[0^{\times 4}, 1]$	CS	.9	.83	.49	.65	.81
$[0^{\times 4}, 1]$	FA^{+}	.9	.78	.49	.65	.85
$[0^{\times 4}, 1]$	$FA^{+/-}$	.9	.76	.49	.65	.89
$[-1.5^{\times 4}, -.5]$	any	0	.31	.12	.19	.34
$[-1.5^{\times 4}, -.5]$	CS	.5	.43	.13	.24	.47
$[-1.5^{\times 4}, -.5]$	FA^{+}	.5	.40	.12	.23	.46
$[-1.5^{\times 4}, -.5]$	$FA^{+/-}$	.5	.40	.13	.24	.54
$[-1.5^{\times 4}, -.5]$	CS	.9	.41	.13	.25	.59
$[-1.5^{\times 4}, -.5]$	FA^{+}	.9	.35	.13	.24	.64
$[-1.5^{\times 4}, -.5]$	$FA^{+/-}$	.9	.31	.13	.24	.68

Web Table 1B. Power comparison of the procedures, all pairwise comparisons case, $n = 100$.

μ'	Covariance	$E(\rho^2)$	Power			
			Disc, Ex	Holm, Ex	Holm, McN	Boot, (5)
$[0^{\times 4}, .4]$	any	0	.27	.23	.27	.31
$[0^{\times 4}, .4]$	CS	.5	.61	.55	.61	.66
$[0^{\times 4}, .4]$	FA^{+}	.5	.65	.59	.64	.68
$[0^{\times 4}, .4]$	$FA^{+/-}$	.5	.82	.77	.80	.82
$[0^{\times 4}, .4]$	CS	.9	.96	.93	.95	.95
$[0^{\times 4}, .4]$	FA^{+}	.9	.95	.91	.93	.94
$[0^{\times 4}, .4]$	$FA^{+/-}$	.9	.97	.95	.96	.97
$[-1.96^{\times 4}, -1.28]$	any	0	.34	.17	.24	.27
$[-1.96^{\times 4}, -1.28]$	CS	.5	.53	.26	.36	.39
$[-1.96^{\times 4}, -1.28]$	FA^{+}	.5	.53	.27	.37	.41
$[-1.96^{\times 4}, -1.28]$	$FA^{+/-}$	.5	.59	.31	.43	.51
$[-1.96^{\times 4}, -1.28]$	CS	.9	.66	.34	.49	.57
$[-1.96^{\times 4}, -1.28]$	FA^{+}	.9	.59	.33	.47	.62
$[-1.96^{\times 4}, -1.28]$	$FA^{+/-}$	.9	.57	.34	.48	.69

Web Table 2A: Power comparison of the procedures, pairwise comparisons against a control case, $n = 25$.

			Power			
μ'	Covariance	$E(\rho^2)$	Disc, Ex	Holm, Ex	Holm, McN	Boot, (5)
$[0^{\times 10}, 1]$	any	0	.34	.25	.33	.41
$[0^{\times 10}, 1]$	CS	.5	.67	.45	.58	.69
$[0^{\times 10}, 1]$	FA^{+}	.5	.66	.42	.56	.69
$[0^{\times 10}, 1]$	$FA^{+/-}$	.5	.77	.46	.61	.78
$[0^{\times 10}, 1]$	CS	.9	.89	.49	.65	.87
$[0^{\times 10}, 1]$	FA^{+}	.9	.86	.48	.64	.88
$[0^{\times 10}, 1]$	$FA^{+/-}$	.9	.88	.49	.65	.92
$[-1.5^{\times 10}, -.5]$	any	0	.41	.11	.18	.34
$[-1.5^{\times 10}, -.5]$	CS	.5	.54	.14	.24	.60
$[-1.5^{\times 10}, -.5]$	FA^{+}	.5	.53	.13	.23	.57
$[-1.5^{\times 10}, -.5]$	$FA^{+/-}$	.5	.56	.13	.23	.65
$[-1.5^{\times 10}, -.5]$	CS	.9	.58	.13	.25	.75
$[-1.5^{\times 10}, -.5]$	FA^{+}	.9	.58	.12	.23	.72
$[-1.5^{\times 10}, -.5]$	$FA^{+/-}$	.9	.58	.12	.23	.74

Web Table 2B. (Identical to Table 3 of the original paper). Power comparison of the procedures, pairwise comparisons against a control case, $n = 100$.

			Power			
μ'	Covariance	$E(\rho^2)$	Disc, Ex	Holm, Ex	Holm, McN	Boot, (5)
$[0^{\times 10}, .4]$	any	0	.26	.23	.26	.30
$[0^{\times 10}, .4]$	CS	.5	.59	.52	.58	.64
$[0^{\times 10}, .4]$	FA^{+}	.5	.63	.57	.62	.66
$[0^{\times 10}, .4]$	$FA^{+/-}$	.5	.80	.74	.78	.80
$[0^{\times 10}, .4]$	CS	.9	.96	.92	.95	.94
$[0^{\times 10}, .4]$	FA^{+}	.9	.94	.90	.92	.94
$[0^{\times 10}, .4]$	$FA^{+/-}$	.9	.97	.95	.96	.97
$[-1.96^{\times 10}, -1.28]$	any	0	.40	.17	.24	.23
$[-1.96^{\times 10}, -1.28]$	CS	.5	.60	.26	.36	.41
$[-1.96^{\times 10}, -1.28]$	FA^{+}	.5	.59	.26	.36	.44
$[-1.96^{\times 10}, -1.28]$	$FA^{+/-}$	.5	.69	.30	.43	.57
$[-1.96^{\times 10}, -1.28]$	CS	.9	.77	.34	.48	.72
$[-1.96^{\times 10}, -1.28]$	FA^{+}	.9	.75	.33	.47	.71
$[-1.96^{\times 10}, -1.28]$	$FA^{+/-}$	.9	.75	.33	.47	.74