Appendix

Proof of Theorem 1

Let the adaptation strategy given in Section 3 be denoted $g_{STANDARD}(\mathbf{X_1})$, where $\mathbf{X_1}$ is the vector of first stage data. We proceed in four steps.

1) If $\delta_{L_1} = 0, \delta_{L_2} = 0, \ldots, \delta_{L_K} = 0$ then $P(FWE) = \alpha$ by the definition of y_{cutoff}. If $\delta_{L_1} \leq 0, \delta_{L_2} \leq 0, \ldots, \delta_{L_K} \leq 0$ then by the location family representation of the multivariate normal distribution, the results of this experiment can be represented as those of that with $\delta_{L_1} = 0, \delta_{L_2} = 0, \ldots, \delta_{L_K} = 0$, with δ_{L_1} subtracted from each response value on L_1, δ_{L_2} subtracted from each response value on L_2, etc. Such subtraction cannot increase $P(FWE)$ by definition of the adaptation strategy and test procedure. Therefore FWER is controlled for $\delta_{L_1} \leq 0, \delta_{L_2} \leq 0, \ldots, \delta_{L_K} \leq 0$.

2) Let $\delta_{L_1} > 0, \delta_{L_2} = 0, \ldots, \delta_{L_K} = 0$. Assume as a worst case that a false rejection could also occur with respect to modification $l_{1,1}$, i.e. assume $\delta_{l_{1,1}} = 0$. $FWE \subseteq \{Reject\ H_{L_2} \cup Reject\ H_{L_3} \cup \ldots \cup Reject\ H_{L_K}\} \cup \{Reject\ H_{L_1} \cap Reject\ H_{l_{1,1}}\}$. Then $P(FWE) \leq P(Reject\ H_{L_2} \cup Reject\ H_{L_3} \cup \ldots \cup Reject\ H_{L_K}) + P(Reject\ H_{L_1} \cap Reject\ H_{l_{1,1}}) \leq \alpha - \alpha_1 + \alpha_1 = \alpha$ by the definition of α_1 and since

$P_{\delta_{L_1}>0, \delta_{L_2}=0, \ldots, \delta_{L_K}=0}(Reject\ H_{L_2} \cup Reject\ H_{L_3} \cup \ldots \cup Reject\ H_{L_K}) \leq$

$P_{\delta_{L_1}=-\infty, \delta_{L_2}=0, \ldots, \delta_{L_K}=0}(Reject\ H_{L_2} \cup Reject\ H_{L_3} \cup \ldots \cup Reject\ H_{L_K}) = \alpha - \alpha_1$ by the definition of the procedure. Also, by the definition of the procedure, the FWER is less if $\delta_{l_{1,1}} \neq$

0. The location family representation of the multivariate normal distribution also allows extension to the case where $\delta_{L_1} \geq 0, \delta_{L_2} \leq 0, \ldots, \delta_{L_K} \leq 0$, as in step 1.

3) Let $\delta_{L_1} > 0, \delta_{L_2} > 0, \delta_{L_3} = 0, \ldots, \delta_{L_K} = 0$. Assume as a worst case that a false rejection could also occur with respect to modification $l_{1,1}$ or $l_{2,1}$ i.e. assume $\delta_{l_{1,1}} = \delta_{l_{2,1}} = 0$. $FWE \subseteq \{Reject\ H_{L_3} \cup Reject\ H_{L_4} \cup \ldots \cup Reject\ H_{L_K}\} \cup \{Reject\ H_{L_1} \cap Reject\ H_{l_1}\} \cup \{Reject\ H_{L_2} \cap Reject\ H_{l_2}\}$. We argue that the probability of this event is less than or equal to α.

First, $P_{\delta_{L_1}>0,\delta_{L_2}>0,\delta_{L_3}=0,\ldots,\delta_{L_K}=0}\ (Reject\ H_{L_3} \cup Reject\ H_{L_4} \cup \ldots \cup Reject\ H_{L_K}) \leq$

$P_{\delta_{L_1}=-\infty,\delta_{L_2}=0,\ldots,\delta_{L_K}=0}\ (Reject\ H_{L_2} \cup Reject\ H_{L_3} \cup \ldots \cup Reject\ H_{L_K}) = \alpha - \alpha_1$ from the definition of α_1.

Second,

$P_{\delta_{L_1}>0,\delta_{L_2}>0,\delta_{L_3}=0,\ldots,\delta_{L_K}=0}\ (\{Reject\ H_{L_1} \cap Reject\ H_{l_1}\} \cup \{Reject\ H_{L_2} \cap Reject\ H_{l_2}\}) =$

$P_{\delta_{L_1}>0,\delta_{L_2}>0,\delta_{L_3}=0,\ldots,\delta_{L_K}=0}\ (\{Reject\ H_{L_1} \cap Reject\ H_{l_1}\}|$

$L_1\ has\ largest\ observed\ first\ stage\ effect\ )*$

$P(L_1\ has\ largest\ observed\ first\ stage\ effect) +$

$P_{\delta_{L_1}>0,\delta_{L_2}>0,\delta_{L_3}=0,\ldots,\delta_{L_K}=0}\ (\{Reject\ H_{L_2} \cap Reject\ H_{l_2}\}|$

$L_2\ has\ largest\ observed\ first\ stage\ effect\ )*$

$P(L_2\ has\ largest\ observed\ first\ stage\ effect) \leq$

$\alpha_1 * \{P_{\delta_{L_1}>0,\delta_{L_2}>0,\delta_{L_3}=0,\ldots,\delta_{L_K}=0}\ (L_1\ has\ largest\ observed\ first\ stage\ effect) +$

$P_{\delta_{L_1}>0,\delta_{L_2}>0,\delta_{L_3}=0,\ldots,\delta_{L_K}=0}\ (L_2\ has\ largest\ observed\ first\ stage\ effect)\}$

$\leq \alpha_1$.

By the definition of the procedure, the FWER is less if $\delta_{l_{1,1}} \neq 0$ or $\delta_{l_{2,1}} \neq 0$. The location family representation of the multivariate normal distribution allows us to extend this to the case $\delta_{L_1} \geq 0, \delta_{L_2} \leq 0, \ldots, \delta_{L_K} \leq 0$, as in step 1.

4) Similar extensions apply to cases through $\delta_{L_1} > 0, \delta_{L_2} > 0, \ldots, \delta_{L_{K-1}} > 0, \delta_{L_K} = 0$. Finally, by construction, if $\delta_{L_1} > 0, \delta_{L_2} > 0, \ldots, \delta_{L_K} > 0$ the maximum possible FWER is $\alpha_1 < \alpha$.

Proof of Theorem 2

Notation: Let $\bar{Y}_{K,i}$ denote the estimated effect (sample mean minus control sample mean) on main dose K in stage i. L^* will denote the main dose selected at the interim analysis. $\bar{Y}_{L^*,overall}$ will denote the overall mean effect on L^*, combining data from both stages. The subscript NULL will refer to quantities calculated under the global null hypothesis, H_{NULL}. Subscripts such as NULL* will refer to modifications of the global null hypothesis as explained further below.

We use $\mathbf{X_1}$ to denote the first stage data. In general it may include all values on primary endpoints as well as additional secondary endpoints. $\mathbf{X_2}$ will denote all second stage data in general including second stage data on the selected main dose, control, and any added modification(or modifications) as well as the counterfactual data on all other main doses and modifications that could have been included. Adaptation rules, denoted for instance by $g(\mathbf{X_1})$ will be functions of the first stage data that determine the selection and promotion decisions applied to the second stage.

For convenience, we will sometimes denote probabilities of outcomes with explicit reference to the adaptation rule in use, with this rule indicated inside the probability statement following a semicolon, eg. $P_{NULL^*}\{\bar{Y}_{L^*,1} > c_1 \cap \bar{Y}_{L^*,overall} > y_{cutoff}; g_{STANDARD}(\mathbf{X_1})\}$.

Let y_{cutoff} be chosen such that $P_{NULL}\left(\bar{Y}_{L^*,1} > c_1 \cap \bar{Y}_{L^*,overall} > y_{cutoff}\right) = \alpha$ under the previously given standard exploration strategy, denoted $g_{STANDARD}(\mathbf{X_1})$. For given c_1, this may be calculated as the α upper critical value of the distribution of $I(\bar{Y}_{L^*,1} > c_1)\bar{Y}_{L^*,overall}$ under H_{NULL}. Let $\alpha^* = P_{NULL^*}(\bar{Y}_{L^*,1} > c_1 \cap \bar{Y}_{L^*,overall} > y_{cutoff})$ under $g_{STANDARD}(\mathbf{X_1})$ where $NULL^*$ is the configuration with $\delta_{L_1} = -\infty, \delta_{L_2} = 0, \ldots, \delta_{L_K} = 0$. Equivalently, α^* may be calculated as $P_{NULL}\left(\bar{Y}_{L^*,1} > c_1 \cap \bar{Y}_{L^*,overall} > y_{cutoff}\right)$ under an exploration strategy $g_{L_2,\ldots,L_K}(\mathbf{X_1})$ that selects the dose with best first stage performance among $L_2, \ldots, L_K$. Define $\alpha_1 = \alpha - \alpha^*$. The testing of any added modification (modifications) of the selected main dose will be at (overall) level α_1 following rejection of the corresponding main dose.

To show strong familywise error rate control over a broad class of selection rules we proceed through four steps. It is assumed that in allowable selection rules the early stopping threshold c_1 must be observed and that the threshold y_{cutoff} may not be changed; however the selection decision and addition of modifications are allowed to depart from $g_{STANDARD}(\mathbf{X_1})$. We assume that the actual adaptation rule used in practice can be represented as some measurable function, $g_{TRUE}(\mathbf{X_1})$, of the first stage data; if desired the domain of this function can also be enlarged to include other data external to the experiment which is available at the interim analysis.

1) For $\delta_{L_1} = 0, \delta_{L_2} = 0, \ldots, \delta_{L_K} = 0$, strong familywise error rate control at level α using $g_{STANDARD}(\mathbf{X_1})$ follows by the definition of y_{cutoff}. Additionally, the strong FWER control holds under $\delta_{L_1} \leq 0, \delta_{L_2} \leq 0, \ldots, \delta_{L_K} \leq 0$ by the location-family property of the normal distributions: subtraction of a constant equal to δ_{L_K} from each $\overline{Y}_{K,i}$ can only lead to decreased rejection of null hypotheses and decreased FWER. Strong familywise error rate control for this parameter configuration holds as well with selection strategies that may choose a dose at interim other than that with best observed first stage performance. This can be thought of as equivalent to the addition of a (nonpositive) random variable equal to the difference between the observed first stage effect on the promoted main dose and the best observed first stage effect and, given unchanged c_1 and y_{cutoff} can only lead to decreased or unchanged FWER. Added leaves may only be tested after (incorrect) rejection of their associated main doses such that their presence or absence or use or disuse does not affect the FWER.

2) Consider a configuration $\boldsymbol{\delta_{L_1}}$ such that $\delta_{L_1} > 0, \delta_{L_2} = 0, \ldots, \delta_{L_K} = 0, \delta_{l_1} = 0$, and the adaptation rule $g_{PREVISION}(\mathbf{X_1}, \mathbf{X_2})$ with 'prevision' in order to maximize the occurrence of familywise errors, as follows. Let $g_{PREVISION}(\mathbf{X_1}, \mathbf{X_2})$ select $\underset{L \in L_2, \ldots, L_K}{ARGMAX}(\overline{Y}_{L,1})$ if it would be rejected in the overall experiment (incorrectly) and otherwise select L_1. Let $g_{PREVISION}(\mathbf{X_1}, \mathbf{X_2})$ also add the leaves on the selected main dose. Clearly, $g_{PREVISION}(\mathbf{X_1}, \mathbf{X_2})$ is impossible in practice since the interim adaptation depends on first stage data only. We note also that, due to the test statistics' construction, addition of the modifications does not change the null distribution of the test statistics related to any main dose.

To first show that $g_{PREVISION}(X_1, X_2)$ leads to FWER ≤ α under $\boldsymbol{\delta_{L_1}}$, consider that

$$P_{\boldsymbol{\delta}_{L_1}}(FWE; g_{PREVISION}(\mathbf{X_1}, \mathbf{X_2})) \leq$$

$$P_{\boldsymbol{\delta}_{L_1}}\left\{\bar{Y}_{L^*,1} > c_1 \bigcap \bar{Y}_{L^*,overall} > y_{cutoff};\ g_{L_2,\ldots,L_K}(\mathbf{X_1})\right\} + \alpha_1 =$$

$$P_{NULL^*}\{\bar{Y}_{L^*,1} > c_1 \cap \bar{Y}_{L^*,overall} > y_{cutoff};\ g_{STANDARD}(\mathbf{X_1})\} + \alpha_1 \leq \alpha^* + \alpha_1 = \alpha.$$

(i)

In the second line, the first term represents a false rejection among main doses $L_2, \ldots, L_K$ and the second term is an upper bound on the probability of a false rejection originating from a tested modification. By the definition of $g_{PREVISION}(\mathbf{X_1}, \mathbf{X_2})$ and the configuration $NULL^*$, this is equal to $P_{NULL^*}\{\bar{Y}_{L^*,1} > c_1 \cap \bar{Y}_{L^*,overall} > y_{cutoff};\ g_{STANDARD}(\mathbf{X_1})\} + \alpha_1 \leq \alpha^* + \alpha_1 = \alpha$, as in the third line.

Also, the increase of the value of any modification's effect above zero cannot cause an increase in the FWER. By the argument used in Step 1, strong FWER control extends to the case of $\delta_{L_2} \leq 0, \ldots, \delta_{L_K} \leq 0$. Intuitively, this follows from the location family property of normal distributions and the consideration that downward shift of the values from any $\bar{Y}_{K,i}$ can only lead to unchanged or decreased rejection of the associated null hypotheses and unchanged or decreased FWER under $g_{PREVISION}(\mathbf{X_1}, \mathbf{X_2})$ or $g_{STANDARD}(\mathbf{X_1})$ given that c_1 and y_{cutoff} are left unchanged.

3) The upper bound of α on the FWER under $\boldsymbol{\delta}_{\boldsymbol{L_1}}$ is next shown to hold over different adaptation rules. Given that the actual adaptation rule followed is $g_{ACTUAL}(\mathbf{X_1})$, we also

consider the rule $g_{ACTUAL\ MAX}(\mathbf{X_1})$, similarly represented as measurable function, that selects L_1 when $g_{ACTUAL}(\mathbf{X_1})$ selects L_1, and selects $\underset{L\in L_2,\ldots,L_K}{ARGMAX}(\bar{Y}_{L,1})$ whenever $g_{ACTUAL}(\mathbf{X_1})$ selects any of $L_2, \ldots, L_K$. Let the rule $g_{ACTUAL\ MAX}(\mathbf{X_1})$ also add the leaves of its chosen main dose. We show

$$P_{\boldsymbol{\delta}_{L_1}}\{FWE; g_{ACTUAL}(\mathbf{X_1})\} \leq$$

$$P_{\boldsymbol{\delta}_{L_1}}\{FWE; g_{ACTUAL\ MAX}(\mathbf{X_1})\} \leq$$

$$P_{\boldsymbol{\delta}_{L_1}}\{FWE; g_{PREVISION}(\mathbf{X_1}, \mathbf{X_2})\}.$$

(ii)

The first inequality follows because the result of the first stage selection of $g_{ACTUAL\ MAX}(\mathbf{X_1})$, considered as a data vector, can be written as that of $g_{ACTUAL}(\mathbf{X}_1)$ plus a random vector whose components are all nonnegative. Thus the probability of a false rejection, can only increase or remain unchanged; that $g_{ACTUAL\ MAX}(\mathbf{X_1})$ enforces addition of modifications can only cause the probability of FWE to increase or remain unchanged.

The second inequality follows from a partition of the sample space as follows.

A_1: $g_{PREVISION}(\mathbf{X_1}, \mathbf{X_2})$ and $g_{ACTUAL\ MAX}(\mathbf{X_1})$ both select L_1

A_2: $g_{PREVISION}(\mathbf{X_1}, \mathbf{X_2})$ selects L_1, $g_{ACTUAL\ MAX}(\mathbf{X_1})$ selects $\underset{L\in L_2,\ldots,L_K}{ARGMAX}(\bar{Y}_{L,1})$

A_3: $g_{PREVISION}(\mathbf{X_1}, \mathbf{X_2})$ selects $\underset{L\in L_2,\ldots,L_K}{ARGMAX}(\bar{Y}_{L,1})$, $g_{ACTUAL\ MAX}(\mathbf{X_1})$ selects L_1

A_4: Both select $\underset{L\in L_2,\ldots,L_K}{ARGMAX}(\bar{Y}_{L,1})$

Then $P_{\boldsymbol{\delta}_{\boldsymbol{L_1}}}\{FWE;\ g_{PREVISION}(\mathbf{X_1},\boldsymbol{X_2})\} - P_{\boldsymbol{\delta}_{\boldsymbol{L_1}}}\{FWE;\ g_{ACTUAL\ MAX}(\mathbf{X_1})\} \geq$

$[P_{\boldsymbol{\delta}_{\boldsymbol{L_1}}}\{FWE|A_2;\ g_{PREVISION}(\mathbf{X_1},\boldsymbol{X_2})\} - P_{\boldsymbol{\delta}_{\boldsymbol{L_1}}}\{FWE|\ A_2;\ g_{ACTUAL\ MAX}(\mathbf{X_1})\ \}]\ P(A_2) +$

$[P_{\boldsymbol{\delta}_{\boldsymbol{L_1}}}\{FWE|\ A_3;\ g_{PREVISION}(\mathbf{X_1},\mathbf{X_2})\} - P_{\boldsymbol{\delta}_{\boldsymbol{L_1}}}\{FWE|A_3;\ g_{ACTUAL\ MAX}(\mathbf{X_1})\ \}]\ P(A_3)$

≥ 0, since each term is greater than or equal to zero by the definitions of $g_{PREVISION}(\mathbf{X_1},\mathbf{X_2})$ and $g_{ACTUAL\ MAX}$.

4) Now consider the case that $\delta_{L_1} > 0, \delta_{L_2} > 0, \ldots, \delta_{L_K} = 0$. The steps (2) and (3) can be repeated by modifying $\boldsymbol{\delta}_{\boldsymbol{L_1}}$ to $\boldsymbol{\delta}_{\boldsymbol{L_1},\boldsymbol{L_2}}$ with $\delta_{L_1} > 0, \delta_{L_2} > 0, \delta_3 = 0, \ldots, \delta_{L_K} = 0$, $\delta_{l_1} = 0$, $\delta_{l_2} = 0$, $NULL^*$ to the configuration with $\delta_{L_1} = -\infty, \delta_{L_2} = -\infty, \delta_{L_3} = 0, \ldots, \delta_{L_K} = 0$, and modifying $g_{PREVISION}(\mathbf{X_1},\mathbf{X_2})$ to select $\underset{L\in L_3,\ldots,L_K}{ARGMAX}(\bar{Y}_{L,1})$ if it would be rejected in the overall experiment (incorrectly), and otherwise to select among L_1 and L_2 so as to maximize type 1 errors (originating from false declaration of efficacy on a modification after correct declaration on a main dose) . The rule $g_{ACTUAL\ MAX}(\mathbf{X_1})$ is modified to select L_1 or L_2 when $g_{ACTUAL}(\mathbf{X_1})$ selects L_1 or L_2 respectively, and to select $\underset{L\in L_3,\ldots,L_K}{ARGMAX}(\bar{Y}_{L,1})$ whenever $g_{ACTUAL}(\mathbf{X_1})$ selects any of $L_3, \ldots, L_K$. The rule $g_{L_2,\ldots,L_K}(\mathbf{X_1})$ that selects the dose with best observed first stage performance among $L_2, \ldots, L_K$ is modified to $g_{L_3,\ldots,L_K}(\mathbf{X_1})$ which selects the dose with best observed first stage performance among $L_3, \ldots, L_K$. The previous partition is modified analogously based on selection of L_1 or L_2 instead of selection of L_1. The result is to confirm FWER control at level α for $g_{ACTUAL}(\mathbf{X_1})$ where $\delta_{L_1} > 0, \delta_{L_2} > 0, \ldots, \delta_{L_K} = 0$ as above. Equations based on (i) and (ii) above continue to apply.

Similar extensions can be made for all cases up to $\delta_{L_1} > 0, \delta_{L_2} > 0, \ldots, \delta_{L_{K-1}} > 0$.

Finally, by construction, if $\delta_{L_1} > 0, \delta_{L_2} > 0, \ldots, \delta_{L_K} > 0$ the maximum possible FWER is $\alpha_1 < \alpha$.

To supplement the above mathematical proofs, further simulation studies based on the designs presented in Table 1 under different adaptation strategies have yielded the following results on the actual level of FWER control achieved by the procedure given in Section 2. The adaptation strategies included: selection of the main dose with best observed first stage performance (with and without addition of a modification), selection of the main dose with worst observed first stage performance (with and without addition of a modification), and random selection among the first stage doses (with and without addition of a modification) using equal probabilities.

1) The FWER is controlled under the global null hypothesis at the specified level $\alpha = .05$ in these designs. The nominal $\alpha = .05$ is achieved under selection of the main dose with best observed first stage performance and this holds whether addition of modifications (leaves) is allowed or not. The FWER is $\leq \alpha = .05$ under all adaptation strategies when the effect of each main dose is ≤ 0.

2) Under all adaptation strategies that do not permit any addition of modifications, the FWER is controlled at a level less than α =.05 for all parameter configurations outside of the global

null. When addition is allowed, the FWER is unaffected unless an added modification is null (zero mean difference relative to control). In this case, enforcing addition of a modification with null effect increases the FWER but the overall FWER still does not exceed the nominal α =.05 in any case. Indeed, under this adaptation strategy, the given α_1 appears to be mildly conservative. The degree of conservatism in the overall FWER appears to be approximately of order α^2, such that the maximum FWER seen in the cases presented in Table 1 over the different adaptation strategies lies between .05 and .0475.