Biometrical Journal **0** (0) 0, zzz–zzz / DOI: 10.1002/bimj.200100000

Multiple Confidence Intervals for Selected Parameters Adjusted for the False Coverage Rate in Monotone Dose-Response Microarray Experiments

Jianan Peng[*1], **Wei Liu** [2], **Frank Bretz** [3], and **Ziv Shkedy** [4]

[1] Department of Mathematics and Statistics, Acadia University, Wolfville, NS, Canada, B4P 2R6
[2] S3RI and School of Mathematics, University of Southampton, SO17 1BJ, UK
[3] Novartis Pharma AG, Basel, 4002, Switzerland, and Shanghai University of Finance and Economics, People's Republic of China
[4] Center for Statistics, Hasselt University, 3590 Diepenbeek, Belgium

Benjamini and Yekutieli (2005) introduced the concept of the false coverage-statement rate (FCR) to account for selection when the confidence intervals are constructed only for the selected parameters. Dose-response analysis in dose-response microarray experiments is conducted only for genes having monotone dose-response relationship, which is a selection problem. In this paper, we consider multiple confidence intervals for the mean gene expression difference between the highest dose and the control in monotone dose-response microarray experiments for selected parameters adjusted for the FCR. A simulation study is conducted to study the performance of the method proposed. The method is applied to a real dose-response microarray experiment with 16,998 genes for illustration.

Key words: Dose-response study; FCR; FDR; Order restricted inference; Selective Inference.

1 Introduction

Dose-response studies are of central importance in pharmaceutical development. For this purpose, a dose-response experiment is conducted in which several doses are administered to separate groups of experimental units. In most cases, a zero-dose control group, or placebo group, is included which serves as a standard against which the dose groups are to be compared with. That is, a one-way layout setup is considered with observations: $Y_{ij} = \mu_i + \epsilon_{ij}, i = 1, \ldots, k, j = 1, \ldots, n_i$, where ϵ_{ij} are iid normal with mean 0 and variance σ^2. The control group is indexed as 1 and the remaining $k-1$ treatment groups correspond to the $k-1$ increasing dose levels, with n_i subjects randomly assigned to group $i = 1, \ldots, k$. The statistic $S^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2/\nu$ is used as an estimator for σ^2, and it is independent of the sample means $\bar{\mathbf{Y}} = (\bar{Y}_1, \ldots, \bar{Y}_k)$, where $\nu S^2/\sigma^2 \sim \chi^2_\nu$ and $\nu = \sum_{i=1}^{k} n_i - k > 0$. Often, the dose-response is monotonic, that is, $\mu_1 \leq \cdots \leq \mu_k$, and it is assumed that a larger μ_i indicates a better average outcome. The primary goal of a dose-response study is to assess whether there is indeed a dose-response effect, and if a dose-response effect is found, then to quantify the lower bound for $\mu_k - \mu_1$ or to identify the lowest dose level producing a desirable effect over that of the control (Tamhane, Hochberg, and Dunnett, 1996), which can be obtained by using confidence intervals (Hsu and Berger (1999) and Peng et al. (2008)) or combining multiple comparisons and modeling (Bretz et al. (2005)).

Recently, dose-response studies are integrated with microarray experiments; see Hu et al. (2005), Lin et al. (2007) and Lin et al. (2012). In a dose-response microarray experiment, the response is the gene expression at a certain dose level and the aforementioned ANOVA model is considered for each gene. The dose-response curve, similar to the classical dose-response studies, is assumed to be monotone, i.e., the

*Corresponding author: e-mail: jianan.peng@acadiau.ca Phone: 1-902-5851586, Fax: 1-902-5851074

gene activity increases or decreases as the dose level increases. The direction of the relationship is usually unknown in advance. A gene is called differentially expressed (DE) if there is an increasing or decreasing trend with respect to dose levels in the gene expression (Lin et al. (2007)). In the microarray dose-respnose experiment used in Lin et al. (2007) and Lin et al. (2008), the human epidermal squamous carcinoma cell line A431 was grown in Dulbecco's modified agle's medium, supplemented with L-glutamined (20 mM), Gentamycin (5 mg/ml) and 10% fetal bovine serum. The microarray experiment had four experimental conditions with three treatments and one placebo. Each condition had three samples hybridized to individual chips. There were 12 arrays and 16,998 genes in the data set; see Lin et al. (2007) and Lin et al. (2008) for details of preprocessing the data set. Below are the gene expression level figures for the first two genes in the data set (the first three values are for the placebo group). The aim of this microarray dose-response oncology experiment was designed to better understand the biological effects of growth factors in human tumor (Lin et al. (2012)). The first step of the study was to find genes whose expression levels are differentiated among the four dose levels.

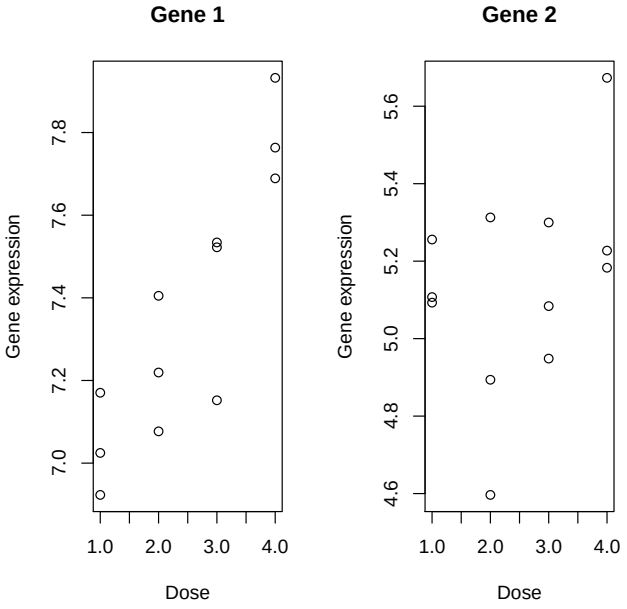

Figure 1 Scatter plots for the gene expression values of the first two genes in the data set

Microarray technology has become an important tool for simultaneously screening thousands of genes for changes in their expression patterns. When simultaneously testing a very large number of hypotheses, the chance for a false positive test result (type I error) increases sharply and an adjustment for multiplicity is needed. In classical dose-response studies, the control of the familywise error rate (FWER), i.e. the probability of at least one type I error, is appropriate. However, the purpose of a microarray experiment is to find genes that are potential candidates for future investigations and the control of false discovery rate (FDR) is often preferred. FDR introudced by Benjamini and Hochberg (1995) is the expected proportion of erroneously rejected null hypotheses among the identified differentially expressed genes. Procedures that control the FDR have been well studied; see Benjamini and Liu (1999) and Benjamini (2010) among many others.

Benjamini and Yekutieli (2005) argued that two types of problems usually arise when considering inference for multiple parameters: simultaneity refers to the need to provide inference that simultaneously

applies to any subset of several parameters and selective inference refers to the need of providing valid inferences for parameters that are selected after viewing the data. In many large scale problems such as dose-response microarray experiments we are often not concerned about simultaneous coverage but do care about the effect of selection on the average marginal properties over the selected genes, since only the findings about the selected ones are of importance (Benjamini (2010)). Benjamini and Yekutieli (2005) coined the false coverage-statement rate (FCR) as the appropriate criterion for capturing the error for confidence intervals constructed for selected parameters and defined in a way that parallels the definition of the FDR. Let R denote the number of confidence intervals constructed, and let V be the number of confidence intervals that do not cover their respective parameters. Then FCR is the expected proportion of intervals failing to cover the parameters among all genes that have been detected as differentially expressed. That is $FCR = E[V/\max(R,1)]$. Tamhane (2005) gave an exact expression for FCR for some special cases.

Dose-response analyses are conducted only for those genes that are found to have either an increasing or a decreasing dose-response relationship. Because the mean dose-response curve for each gene is a multivariate object, dose-response analyses in microarray experiments can be viewed as a selection-adjusted simultaneity problem (Lin et al., 2012).

Lin et al. (2012) modified Benjamini and Yekutieli (2005)'s FCR-adjusted Benjamini-Hochberg (BH) selected confidence intervals (cis) for $\frac{\sum c_i\mu_i}{\sum d_i\mu_i}$, where coefficients $\mathbf{c} = (c_1, \ldots, c_k)$ and $\mathbf{d} = (d_1, \ldots, d_k)$ are contrasts. However, Lin et al. (2012)'s two-sided confidence interval for the ratio $\frac{\sum c_i\mu_i}{\sum d_i\mu_i}$ does not take the monotonicity of the doses into account even though the monotonicity was assumed. Gene expression data are usually log-transformed by one of several data preprocessing steps. Therefore, a ratio statistic is less attractive (Lin et al., 2012). Jung et al. (2011) considered FCR adjusted confidence intervals for two sample microarray data and argued that confidence intervals are helpful to assess genes' biological relevance. However, ANOVA type dose-response microarray experiments are very common in pharmaceutical research and development. ICH guideline E9 "Statistical Principle for Clinical Trials" suggests that estimates of treatment effects should be accompanied by confidence intervals, whenever possible (see www.ich.org). Even though ICH guideline E9 was developed for classical dose-response studies, it is very important to adjust the marginal level of multiple confidence intervals for selected genes in dose-response microarray experiments as FDR control is widely used in microarray experiments. In this paper, we modify Benjamini and Yekutieli (2005)'s FCR-adjusted BH-selected confidence intervals for mean differences such as $\mu_k - \mu_1$ in ANOVA type dose-response microarray experiments. FCR-adjusted confidence intervals for $\mu_k - \mu_1$ are useful to assess the sizes of the improvement of a drug for selected genes.

This paper is organized as follows. In Section 2 we review the construction of the lower confidence bound for $\mu_k - \mu_1$ under the order restriction $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$. In Section 3 we review the selection of DE genes. We propose a new procedure in Section 4. Simulation results are discussed in Section 5. We apply the new procedure to the aforementioned real microarray data set in Section 6. A brief discussion is given in Section 7.

2 The Lower Confidence Bound for $\mu_k - \mu_1$ under Monotonicity

In this section we first review how to test whether there is a dose-response effect by testing $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ versus $H_1 : \mu_1 < \mu_k$ under the monotonicity assumption $\mu_1 \leq \cdots \leq \mu_k$ and then how to construct a lower confidence bound for $\mu_k - \mu_1$. The method in this section also works for constructing a lower confidence bound for $\mu_1 - \mu_k$ under the antitonic assumption $\mu_1 \geq \cdots \geq \mu_k$. Although there are many test statistics available for testing H_0 versus H_1, only a few can provide interval inference. We review the multiple contrast test statistic to construct the lower confidence bound for $\mu_k - \mu_1$ in Peng et al. (2008). The method in this section can be applied for classical montone dose-response studies or for a single prespecified gene in the monotone dose-response microarray experiment discussed in Sections 1 and 4.

2.1 The Likelihood Ratio Test

Under the monotonicity assumption the likelihood ratio test rejects H_0 in favor of H_1 for large values of

$$S_{01} = \sum_{i=1}^{k} n_i(\mu_i^* - \hat{\mu})^2 / \Big\{ \sum_{i=1}^{k} n_i(\bar{Y}_i - \mu_i^*)^2/\nu + S^2 \Big\},$$

where $\hat{\mu} = \sum_{i=1}^{k} n_i \bar{Y}_i / \sum_{i=1}^{k} n_i$ and μ_i^* are the isotonic (maximum likelihood) estimates of μ_i under the monotonicity assumption, which can be computed by the Pool-Adjacent-Violators Algorithm (PAVA) (see Robertson, Wright and Dykstra, 1988). The null distribution of S_{01} under H_0 is given by

$$P(S_{01} > s) = \sum_{j=2}^{k} P(j,k;\mathbf{w}) P\Big\{ F_{j-1,N-j} > \frac{s(N-j)}{\nu(j-1)} \Big\} \tag{1}$$

for any $s > 0$, where $N = \sum_{i=1}^{k} n_i$, $\mathbf{w} = (n_1, \ldots, n_k)$, and $P(j,k;\mathbf{w})$ is the level probability under H_0 that $\boldsymbol{\mu}^*$ takes j distinct values (see Section 2.4 of Robertson, Wright and Dykstra, 1988). The critical value $s_{k,\nu,\alpha}$ of the statistic S_{01} can be found in Robertson, Wright and Dykstra (1988) or obtained by (4).

2.2 The Multiple Contrast Test Statistic T

When $S_{01} > s_{k,\nu,\alpha}$, one rejects H_0 and concludes that there is indeed a dose-response effect, that is, at least one mean response μ_i $(2 \le i \le k)$ is significantly larger than μ_1. However, there is no corresponding lower confidence bounds for $\mu_k - \mu_1$ when $k > 2$. Peng et al. (2008) introduced the following multiple contrast test statistic:

$$T = \max_{\mathbf{c} \in \mathbf{C}} \sum_{i=1}^{k} n_i c_i \bar{Y}_i / S \Big(\sum_{i=1}^{k} n_i c_i^2 \Big)^{1/2}, \tag{2}$$

where $\mathbf{C} = \Big\{ \mathbf{c} = (c_1, \ldots, c_k) : \sum_{i=1}^{k} n_i c_i = 0,\ c_1 \le \cdots \le c_k \Big\}$. It can be proved that the statistic T^2 is asymptotically equivalent to S_{01} and

$$T^2 = \sum_{i=1}^{k} n_i(\mu_i^* - \hat{\mu})^2 / S^2. \tag{3}$$

Let $t_{k,\nu,\alpha}$ be the critical value of T,

$$P_{\boldsymbol{\mu}} \Big\{ \sum_{i=1}^{k} n_i c_i \mu_i \ge \sum_{i=1}^{k} n_i c_i \bar{Y}_i - t_{k,\nu,\alpha} S (\sum_{i=1}^{k} n_i c_i^2)^{1/2},\ \text{for all } \mathbf{c} \in \mathbf{C} \Big\} = 1 - \alpha. \tag{4}$$

The left-hand side of (4) can be rewritten as

$$\begin{aligned} & P_{\boldsymbol{\mu}} \Big\{ \max_{\mathbf{c} \in \mathbf{C}} \sum_{i=1}^{k} n_i c_i (\bar{Y}_i - \mu_i) / S (\sum_{i=1}^{k} n_i c_i^2)^{1/2} \le t_{k,\nu,\alpha} \Big\} \\ = \quad & P_{\mathbf{0}} \Big\{ \max_{c \in \mathbf{C}} \sum_{i=1}^{k} n_i c_i \bar{Y}_i / S (\sum_{i=1}^{k} n_i c_i^2)^{1/2} \le t_{k,\nu,\alpha} \Big\} \\ = \quad & P_{\mathbf{0}} \Big\{ T^2 \le t_{k,\nu,\alpha}^2 \Big\}. \end{aligned}$$

The null distribution of T under H_0 is given by

$$P(T \geq t) = \sum_{j=2}^{k} P(j,k;\mathbf{w}) P\Big\{F_{j-1,\nu} \geq \frac{t^2}{j-1}\Big\} \tag{5}$$

for any $t > 0$. The critical value $t_{k,\nu,\alpha}$ is the solution when one equates (5) to α. In case of equal sample sizes, one can obtain $t_{k,\nu,\alpha}$ by applying Table A.10 in Robertson et al. (1988) to (5); if further the variance σ^2 is known, then $t_{k,\nu,\alpha}$ is the square root of the corresponding critical value in Table A. 4 of Robertson et al. (1988). For unequal sample sizes case, one can use the method in Miwa, Hayter, Liu (2000) to obtain the critical values of T; see also Genz and Bretz (2009).

2.3 The Lower Confidence Bound for $\mu_k - \mu_1$

According to (4), the $100(1-\alpha)\%$ one-sided simultaneous confidence bounds for any contrast $\sum_{i=1}^{k} n_i c_i \mu_i$ with $c_1 \leq \cdots \leq c_k$, is given by

$$l\Big(\sum_{i=1}^{k} n_i c_i \mu_i\Big) = \sum_{i=1}^{k} n_i c_i \bar{Y}_i - t_{k,\nu,\alpha} S(\sum_{i=1}^{k} n_i c_i^2)^{1/2}.$$

However, we only focus on the lower confidence bound $\mu_k - \mu_1$ because this quantity may be useful to the experimenter in assessing the actual treatment effect difference between the largest dose and the control. Let $\mathcal{K} = \Big\{\mathbf{c} : \mathbf{c} \in \mathbf{C},\ \Sigma_{i=1}^{k} n_i c_i \mu_i \leq \mu_k - \mu_1,\ \text{for all}\ \boldsymbol{\mu} \in \Omega\Big\}$, where $\Omega = \{\boldsymbol{\mu} : \mu_1 \leq \cdots \leq \mu_k\}$. The largest lower confidence bound for $\mu_k - \mu_1$ is given by

$$L(\mu_k - \mu_1) = \max_{\mathbf{c} \in \mathcal{K}} l\Big(\sum_{i=1}^{k} n_i c_i \mu_i\Big). \tag{6}$$

Peng et al. (2008) proved that the optimization solution to (6) is equivalent to the solution of the optimization problem:

$$\max\{\sum_{i=1}^{k} n_i c_i \mu_i^* - t_{k,\nu,\alpha} S(\sum_{i=1}^{k} n_i c_i^2)^{1/2}\}$$

subject to $\mathbf{c} \in \mathbf{C}$ and $\sum_{i=j}^{k} n_i c_i \leq 1, j = 2, \ldots, k$.

The optimized solution $\mathbf{c}^o$ can be obtained iteratively in a few steps by using the algorithm in Peng et al. (2008). In Section 4 below the largest lower confidence bound for $\mu_{gk} - \mu_{g1}$ for a selected gene g will be denoted as $L(\mu_{gk} - \mu_{g1})$.

3 Selection of DE Genes in Monotone Dose-Response Microarray Experiments

In this section we consider how to obtain DE genes in a monotone dose-response microarray experiments. To identify genes that are differentially expressed between several treatment conditions, employing an ANOVA type of model is one option. In this section, we follow the gene-specific linear model similar to that in Lin et al. (2008). Let $Y_{gij} = \mu_{gi} + \epsilon_{gij}$ be the gth gene expression ($g = 1, \ldots, m$) on array $j = 1, \ldots, n_{gi}$ in treatment group $i = 1, \ldots, k$, where μ_{gi} is the mean expression level for treatment i for gene g and $\epsilon_{gij} \sim N(0, \sigma_g^2)$ independently. In the simulation studies in Section 5 we will allow dependence among genes. For a single gene, it is a typical ANOVA with equal variances among treatments. Therefore, the method in Section 1 can be applied for each gene, but we need to adjust for multiplicity.

For each gene we use the likelihood ratio test (or the multiple contrast test statistic in Section 2) for testing $H_{g0} : \mu_{g1} = \mu_{g2} = \cdots = \mu_{gk}$ against $H_{g1} : \mu_{g1} < \mu_{gk}$ under the monotonicity assumption $\mu_{g1} \leq \cdots \leq \mu_{gk}$ to obtain the raw p-values $p_g, g = 1, \ldots, m$. The case of decreasing profiles is treated similarly. These raw p-values are then adjusted to control the FDR at a pre-specified level, say 5%. The most commonly used FDR adjusting procedure is the BH procedure in Benjamini and Hochberg (1995). For a given FDR level q, the BH procedure sorts the p-values of the m hypotheses, $p_{(1)} \leq \cdots \leq p_{(m)}$. Then the ordered p-value $p_{(g)}$ is compared with the critical value $\frac{q \times g}{m}$. Let $K = \max\{g : p_{(g)} \leq \frac{q \times g}{m}\}$ and reject the corresponding K hypotheses if such a K exists. Otherwise do not reject any null hypotheses. The adjusted p values are $\tilde{p}_{(g)} = \min_{h=g,\ldots,m}[\min(\frac{m}{h} p_{(h)}, 1)]$. Benjamini and Yekutieli (2001) proved that the BH procedure controls the FDR when the test statistics are positively dependent in the sense of exhibiting positive regression dependency on the subset of statistics corresponding to true null hypotheses, called PRDS in their paper (Benjamini and Yekutieli (2005) gave the PRDS definition again on page 77). Sarkar (2002) extended Benjamini and Yekutieli (2001)'s work to a generalized stepwise procedure. Distributions satisfying PRDS include multivariate normal distributions that arise in many-to-one comparisons of means with one-sided alternatives and known variances in clinical trials, among others, see Benjamini and Yekutieli (2001) and Sarkar (2002). Yekutieli (2008) developed a modification of the BH procedure to control FDR for testing non-positive dependent test statistics. Benjamini and Yekutieli (2001) proposed the BY procedure to allow test statistics to have an arbitrary dependence structure but it is more conservative than the BH procedure. Its adjusted p values are $\tilde{p}_{(g)} = \min_{h=g,\ldots,m}[\min(\frac{m \Sigma_{l=1}^{m} \frac{1}{l}}{h} p_{(h)}, 1)]$. Note that the rejection of the null hypothesis $H_{g0} : \mu_{g1} = \mu_{g2} = \cdots = \mu_{gk}$ does not indicate the magnitude by which gene expression increases or decreases (Lin et al. (2012)). The confidence lower bounds for $\mu_{gk} - \mu_{g1}$ and $\mu_{g1} - \mu_{gk}$ among the selected DE genes may be useful for a researcher to select genes whose expressions are significantly larger than a threshold.

4 FCR-Adjusted BH Selected CIs Procedure for Monotone Dose-Response Microarray Experiments

In monotone dose-response microarray experiments, the dose-reponse (gene activity) increases or decreases as the dose level increases. For each gene g these two monotone trends cannot hold simultaneously. However, the direction of the dose-response effects is unknown in advance. Therefore, the parameters that we are interested are selected from the union of $\{\mu_{gk} - \mu_{g1}, g = 1, \ldots, m\} \cup \{\mu_{g1} - \mu_{gk}, g = 1, \ldots, m\}$. We modify the procedure of constructing FCR-adjusted BH-selected confidence intervals in Benjamini and Yekutieli (2005) by introducing the following four steps:

1. Use the likelihood ratio test (or the multiple contrast test statistic in Section 2) for $H_{g0} : \mu_{g1} = \mu_{g2} = \cdots = \mu_{gk}$ versus $H_{g1} : \mu_{g1} \leq \cdots \leq \mu_{gk}$ and H_{g0} versus $H'_{g1} : \mu_{g1} \geq \cdots \geq \mu_{gk}$ with at least one strict inequality for each gene to have m raw p-values and sort them as $p_{(1)} \leq \cdots \leq p_{(m)}$ and $p'_{(1)} \leq \cdots \leq p'_{(m)}$ for each trend. The test statistics are denoted as **T=** $\{T_g, 1 \leq g \leq m\}$ and **T′=** $\{T'_g, 1 \leq g \leq m\}$.

2. Calculate $R_{CI} = R + R'$, where $R = \max\{g : p_{(g)} \leq \frac{g \times q/2}{m}\}$, $R' = \max\{g : p'_{(g)} \leq \frac{g \times q/2}{m}\}$, and q is the required FDR level.

3. Select the R_{CI} parameters, for which $p_{(g)} \leq \frac{Rq/2}{m}$ (or $p'_{(g)} \leq \frac{R'q/2}{m}$) , corresponding to the rejected hypotheses.

4. Construct a $1 - \frac{R_{CI} q/2}{m}$ confidence lower bound $L(\mu_{gk} - \mu_{g1})$ for $\mu_{gk} - \mu_{g1}$ (or $L(\mu_{g1} - \mu_{gk})$ for $\mu_{g1} - \mu_{gk}$) for the selected gene g based on the method in Peng et al. (2008) for each of the R_{CI} selected DE genes.

If one knows the direction of the monotone dose-response effects in advance such as the simulaton studies in Section 5, one should use the original FCR-adjusted BH selected cis in Benjamini and Yetutieli (2005) instead of the modified one above. According to Theorem 1 in Benjamini and Yekutieli (2005) if the genes are independent from each other, that is, the test statistics $\mathbf{T}$ and $\mathbf{T}'$ among m genes are independent, then the above procedure ensures that FCR $\leq q$. In order to prove that the above procedure for selective CIs is concordant, we have the following theorem (suggested by one of the referees) first. For simplicity, we only state for the increasing trend case.

Theorem 4.1 *Assume that the monotonicity assumption $\mu_{g1} \leq \cdots \leq \mu_{gk}$ holds for $g = 1, \ldots, m$. If gene g is selected by either the Bonferroni procedure or the FCR-adjusted BH selected procedure given in Definition 1 on page 73 of Benjamini and Yetutieli (2005) for monotone dose-response microarray experiments above, then the lower bound $L(\mu_{gk} - \mu_{g1}) > 0$.*

Proof. Note that we have the assumption of increasing trends in the theorem, we use the original FCR-adjusted BH selected cis in Benjamini and Yetutieli (2005) instead of the modified one above. If a gene is selected by either procedure in the theorem, then its corresponding p-value is less than or equal to the significance level $\frac{q}{m}$ or $\frac{R_{CI}q}{m}$. Following Theorem 2.1 in Peng et al. (2008), the lower bound $L(\mu_{gk} - \mu_{g1}) > 0$.

Proposition 4.2 *Both the above procedure for selective CIs and the one-sided confidence intervals in Section 2 are concordant.*

Proof. We prove that $\{T_g : \theta_g \notin CI_g(\alpha)\}$ is concordant in Definition 5 of Benjamini and Yekutieli (2005), where $\theta_g = \mu_{gk} - \mu_{g1}$ and $CI_g(\alpha) = (L^\alpha(\mu_{gk} - \mu_{g1}), +\infty)$, where $L^\alpha(\mu_{gk} - \mu_{g1})$ is the lower bound in Theorem 4.1 with confidence level $100(1-\alpha)\%$. The cases for $\{T'_g : \mu_{g1} - \mu_{gk} \notin (L^\alpha(\mu_{g1} - \mu_{gk}), +\infty)\}$ and $\{\mathbf{T}^{(g)} : k \leq R_{min}(\mathbf{T}^{(g)})\}$ can be done similarly. For $\alpha \leq \alpha'$, $\{T_g : \theta_g \notin CI_g(\alpha)\} = \{T_g : \theta_g \leq L^\alpha(\mu_{gk} - \mu_{g1})\}$. Note that $L^\alpha(\mu_{gk} - \mu_{g1}) \leq L^{\alpha'}(\mu_{gk} - \mu_{g1})$. Therefore, $\{T_g : \theta_g \notin CI_g(\alpha)\} \subseteq \{T_g : \theta_g \notin CI_g(\alpha')\}$.

Based on Theorem 3 in Benjamini and Yekutieli (2005) and Proposition 4.2, we have the following result.

Theorem 4.3 *If the components of $\mathbf{T}$ and $\mathbf{T}'$ are PRDS, the FCR-adjusted selective CIs above enjoy $FCR \leq q$.*

Note that Steps 1 and 2 are modified from the FDR controlling testing procedure of BH. We can prove that it controls the FDR by using the method in Section 6 in Benjamini and Yekutieli (2005) (suggested by the referee).

Theorem 4.4 *If the components of $\mathbf{T}$ and $\mathbf{T}'$ are PRDS, then the FDR in Steps 1 and 2 of the above procedure is controlled at level q.*

Proof. As all $g \in S(\mathbf{T})$, the selection procedure, have $p_g \leq \frac{R \times q/2}{m}$, where p_g is the corresponding p-value. Applying the FCR in Step 4 implies that all one-sided intervals constructed are $(L^{\frac{R_{CI}q/2}{m}}(\mu_{gk} - \mu_{g1}), +\infty)$. Accodring to Theorem 4.1, these $L^{\frac{R_{CI}q/2}{m}}(\mu_{gk} - \mu_{g1}) > 0$. Similarly for all $g' \in S(\mathbf{T}')$ have $p'_{g'} \leq \frac{R' \times q/2}{m}$ associated with constructed one-sided intervals $(L^{\frac{R_{CI}q/2}{m}}(\mu_{g'1} - \mu_{g'k}), +\infty)$. Therefore, V_{CI} is the number of $g \in S(\mathbf{T})$ for which $\mu_{gk} - \mu_{g1} = 0 \notin (L^{\frac{R_{CI}q/2}{m}}(\mu_{gk} - \mu_{g1}), +\infty)$ plus the number of $g' \in S(\mathbf{T}')$ for which $\mu_{g'1} - \mu_{g'k} = 0 \notin (L^{\frac{R_{CI}q/2}{m}}(\mu_{g'1} - \mu_{g'k}), +\infty)$. That means the number of true null hypotheses rejected by the modified BH procedure equals V_{CI}. Therefore, $FDR = FCR \leq q$ by Theorem 4.3.

In dose-response microarray experiments researchers are interested in testing whether genes are differently expressed and then doing interval inference for selected genes. The dependence of the test statistics $\mathbf{T}$ and $\mathbf{T}'$ among genes in Theorems 4.3 and 4.4 above is unknown in such applications. It is difficult to

check whether or not the test statistics $\mathbf{T}$ and $\mathbf{T}'$ satisfy the PRDS condition. However, it is not unreasonable to assume that test statistics $\mathbf{T}$ and $\mathbf{T}'$ among genes satisfy the PRDS condition. In Section 5, the simulation studies show that the proposed procedure controls FCR even for test statistics $\mathbf{T}$ among genes do not satisfy the PRDS condition.

The adjusted marginal confidence level for selected parameter increases when the proportion of number of parameters/genes selected R_{CI} over the number of considered parameters m decreases in Step 4. Both Step 1 and Step 4 take into account of the monotonicity of the dose-response microarray experiments.

Note that one can use the Bonferroni procedure as a selection rule, when the adjustment of the confidence level ensures control of the FCR-the confidence level is $1 - \frac{Rq}{m}$, where R is the number of selected parameters. However, it is well known that the Bonferroni procedure is very conservative. If the Bonferroni procedure for selection of parameters and the $1 - \frac{Rq}{m}$ confidence lower bounds for the R selected parameters are constructed, the lower bounds are smaller than those given FCR-adjusted BH selected CIs (see Sections 5 and 6).

For general dependency among the components of $\mathbf{T}$ and $\mathbf{T}'$, we can have the FCR-adjusted BY-selected confidence intervals if we replace q by $q\sum_{j=1}^{2m} 1/j$ in Step 2 and change q to $q/(\sum_{j=1}^{2m} 1/j)$ in Step 4. But it is more conservative than the procedure for constructing FCR-adjsuted BH-selected CIs.

Theorem 4.5 *The FCR-adjusted BY-selective CIs is bounded by* $q\sum_{j=1}^{2m} 1/j$.

Proof: Since the critical value $t_{k,\nu,\alpha} \geq t_{k,\nu,\alpha'}$ *for* $\alpha \leq \alpha'$*, the marginal CIs in Section 2 are monotone. Then using Theorem 4 in Benjamini and Yekutieli (2005) by replacing* m *with* $2m$ *we complete the proof.*

Note that since the marginal CIs considered in the paper are monotone, the result above is also true for any selection procedure including the Bonferroni procedure for any dependency structure of the test statistics.

5 Simulation Studies

In this section, we conduct simulation studies to assess the performance of the proposed procedure. We consider a dose-response microarray experiment with $m = 200$ genes and 4 dose levels. The sample size is $n = 10$ for each dose level. Many assumptions on the dependence structure among genes have been made in microarray studies. However, few if any of these assumptions have been verified. It is widely believed that genes are more likely to be dependent within a group, and that each group is independent of the others. In the simulation studies, we consider 20 independent groups and each group has 10 genes. We consider two types of dependence structure among 10 genes, $\Sigma_{1,\rho}$ and $\Sigma_{2,\rho}$, where

$$\Sigma_{1,\rho,i,j} = \begin{cases} 1 & \text{when } i = j, \\ \rho & \text{when } i \neq j, \end{cases}$$

$$\Sigma_{2,\rho,i,j} = \begin{cases} 1 & \text{when } i = j, \\ \rho & \text{when } i < j \leq 5, \\ -\rho & \text{when } i \leq 5, j > 5, \end{cases}$$

and $0 < \rho < 1$. Many papers use $\Sigma_{1,\rho}$ in their simulation studies (Benjamini et al. (2006) among many others). The dependence structure of $\Sigma_{2,\rho}$ (Storey et al. (2004)), does not satisfy the positive regression dependence condition of Benjamini and Yekutieli (2001). Theorem 3 in Benjamini and Yekutieli (2005) cannot be applied here.

We generate 20 groups of genes with 10 genes in each group. The effects of the $m = 200$ genes at each dose level are generated from a multivariate normal distribution with mean vector $\boldsymbol{\mu}_i = (\mu_{1i}, \ldots, \mu_{mi})^T, i =$

$1, \ldots, 4$ and covariance matrix Σ, where

$$\Sigma = \begin{pmatrix} \sigma_1^2\Sigma_\rho & 0 & \cdots & 0 \\ 0 & \sigma_2^2\Sigma_\rho & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{20}^2\Sigma_\rho \end{pmatrix}_{m\times m},$$

$\Sigma_\rho = \Sigma_{1,\rho}$ or $\Sigma_{2,\rho}$ and $\sigma_r^2 = r, r = 1, \ldots, 20$. For simplicity we only consider the increasing trend $\mu_{g1} \leq \mu_{g2} \leq \mu_{g3} \leq \mu_{g4}, g = 1, 2, \ldots, 200$ in the simulations. In this case, the parameters are selected from $\{\mu_{gk} - \mu_{g1}, g = 1, \ldots, 200\}$. We use the original FCR-adjusted BH-selected cis in Benjamini and Yekutieli (2005) and the Bonferroni selection rule with FDR=0.05 instead of $\frac{0.05}{2}$ as we assume that the dose-response for each gene follows the increase trend in the simulation. We consider $\tau = 0.1$ or 0.2 as the proportion of nonzero values in the mean vector $\boldsymbol{\mu}_i$. The first $(1-\tau)\times m$ components in $\boldsymbol{\mu}_i$ are zero. For $g = m(1-\tau)+1, \ldots, m$, configuration $(10, 20, 30, 40)$ in Tables 1 and 3 below means that $\mu_{g1} = 10, \mu_{g2} = 20, \mu_{g3} = 30, \mu_{g4} = 40$. Similarly for other configurations considered in the simulation. The dose-response functions are of two types, linear ($(10, 20, 30, 40)$) and step ($(10, 10, 10, 40), (10, 10, 40, 40)$, and $(10, 40, 40, 40)$). For each configuration, we used 1000 replications. Table 1 (Table 3) gives the simulated FCRs based on BH and Bonferroni selection methods when the gene expression levels are from the multivariate normal distribution with covariance associated with $\Sigma_{1,\rho}$ ($\Sigma_{2,\rho}$). Both tables show that the FCR clearly falls below 5% for the two types of dependence structure. The FCR increases as τ increases. The FCR for BH selection decreases roughly as ρ increases for covariance associated with $\Sigma_{1,\rho}$ but it is not the case for covariance associated with $\Sigma_{2,\rho}$. The simulated FCR in Tables 1 and 3 indicates that among all DE genes only a very small portion of confidence lower bounds does not cover the parameter $\mu_{g4} - \mu_{g1}$.

We also simulate the mean value and the standard deviation of the obtained lower bounds across the 1000 replications for each configuration in Table 1. The results (standard deviations are in the parentheses) are reported in Table 2. It can be seen that both the mean value and the standard deviation of the lower bounds by the Bonferroni procedure are much smaller than those by the BH procedure for the scenarios considered. It seems that the correlation among genes in a group does not affect the mean values of the lower bounds for both procedures, but the proportion of DE genes does affect the mean values of the lower bounds for both procedures. For the configuration in Table 2, the averages of the lower bounds increase when the proportion of DE genes increases from 10% to 20%. However, it seems that the larger the correlation among genes in a group, the larger the standard deviation of the lower bounds for both procedures. It is interesting that the standard deviations of the lower bounds for the BH procedure decrease but the standard deviations of the lower bounds for the Bonferroni procedure increase when the proportion of DE genes increases from 10% to 20%.

6 Revisit the dose-response microarray data set

In this section, we construct the FCR-adjusted BH selected lower confidence intervals for $\mu_{g4} - \mu_{g1}$ (or $\mu_{g1} - \mu_{g4}$) of DE genes. The DE genes are found by the likelihood ratio test for $H_{g0} : \mu_{g1} = \mu_{g2} = \mu_{g3} = \mu_{g4}$ versus $H_{g1} : \mu_{g1} \leq \cdots \leq \mu_{g4}$ and H_{g0} versus $H'_{g1} : \mu_{g1} \geq \mu_{g2} \geq \mu_{g3} \geq \mu_{g4}$ with at least one strict inequality for the 16998 genes in the data set mentioned in Section 1 at FDR $\frac{0.05}{2}$ for each of the monotone directions with BH-FDR adjustment. We have 1197 DE genes with $H_{g1} : \mu_{g1} \leq \cdots \leq \mu_{g4}$ and 1549 DE genes with $H'_{g1} : \mu_{g1} \geq \cdots \geq \mu_{g4}$. We select those $1197 + 1549 = 2746$ DE genes. The FCR-adjusted $1 - R_{CI} \times 0.025/m = 1 - 2746 \times 0.025/16998 = 99.60\%$ lower confidence bounds for the selected mean differences $\mu_{g4} - \mu_{g1}$ (or $\mu_{g1} - \mu_{g4}$) are constructed by the method in Peng et al. (2008). This is because the parameters are selected from the union of parameters: $\{\mu_{g4} - \mu_{g1}, g = 1, \ldots, m\} \cup \{\mu_{g1} - \mu_{g4}, g = 1, \ldots, m\}$ as we do not know the direction of dose-response effects in advance. It is the same reason for the Bonferroni procedure at level $1 - \frac{0.025}{16998}$ and the unadjusted procedure at level $1 - 0.025$ for each direction.

Table 1 Simulated FCR of FCR-adjusted 95% Confidence Lower Bounds for BH and Bonferroni Level 0.05 Selection for Covariance with $\Sigma_{1,\rho}$.

Configuration	τ	Selection Method	ρ			
			0.0	0.2	0.5	0.9
(10, 20, 30, 40)	0.1	BH	0.00151	0.00117	0.00118	0.00038
		Bon	0.00009	0.00005	0.00000	0.00000
	0.2	BH	0.00354	0.00325	0.00240	0.00185
		Bon	0.00010	0.00005	0.00002	0.00000
(10, 10, 10, 40)	0.1	BH	0.00303	0.00241	0.00221	0.00123
		Bon	0.00019	0.00024	0.00010	0.0000
	0.2	BH	0.00634	0.00598	0.00451	0.00417
		Bon	0.00019	0.00024	0.00012	0.00000
(10, 10, 40, 40)	0.1	BH	0.00432	0.00353	0.00260	0.00157
		Bon	0.00024	0.00024	0.00010	0.00010
	0.2	BH	0.00882	0.00776	0.00679	0.00542
		Bon	0.00020	0.00019	0.00017	0.00007
(10, 40, 40, 40)	0.1	BH	0.00281	0.00219	0.00184	0.00174
		Bon	0.00014	0.00010	0.00000	0.00010
	0.2	BH	0.00655	0.00572	0.00509	0.00460
		Bon	0.00015	0.00010	0.00007	0.00022

The following lower confidence bounds for the selected mean differences $\mu_{g4} - \mu_{g1}$ of selected DE genes for the increasing trend are the snapshot of confidence lower bounds based on different adjustments for the 1197 chosen genes. Only 172 Bonferroni-adjusted confidence lower bounds are positive for the 1197 chosen genes by the BH procedure. Below are the three types of confidence lower bounds for the first 10 genes chosen by the Bonferroni selection method.

- The first 10 Bonferroni-adjusted confidence lower bounds for $\mu_{g4} - \mu_{g1}$ of selected DE genes by the Bonferroni procedure are 1.43670859 0.06452051 4.03226602 0.01434778 2.21203461 1.38753585 0.28916412 0.07269671 0.22573569 0.25780339
- The corresponding 10 FCR-adjusted confidence lower bounds of the selected DE genes by the Bonferroni procedure are 2.0856934 0.4874672 5.1134994 0.9831878 2.9122748 3.7325525 0.7729919 0.4861990 1.8055545 0.6062588
- The corresponding 10 unadjusted confidence lower bounds of the selected DE genes by the Bonferroni procedure are

  2.1793371 0.5484954 5.2695134 1.1244278 3.0169314 4.0727143 0.8428411 0.5536372 2.0335108 0.6586269

We also present the three types of confidence lower bounds above in Figure 2. We can see that the unadjusted marginal 97.5% lower confidence bounds for $\mu_{g4} - \mu_{g1}$ are larger than the FCR-adjusted lower bounds and the FCR-adjusted lower bounds are larger than the Bonferroni-adjusted lower bounds at the confidence level of $1 - 0.025/16998 = 99.99985\%$ as expected. This can also be observed from Figure 3 below.

Figure 3 shows the heights of the confidence lower bounds for $\mu_{g4} - \mu_{g1}$ for the first 150 DE genes by the Bonferroni procedure for the increasing trend in the data set in terms of their heights. The unadjusted

Table 2 Simulated Averages and Standard Deviations of FCR-adjusted 95% Confidence Lower Bounds for BH and Bonferroni Level 0.05 Selection for Covariance with $\Sigma_{1,\rho}$.

Configuration	τ	Selection Method	ρ			
			0.0	0.2	0.5	0.9
(10, 20, 30, 40)	0.1	BH	22.55 (1.16)	22.58 (1.35)	22.65 (1.67)	22.76 (2.43)
		Bon	2.15 (0.05)	2.15 (0.08)	2.14 (0.11)	2.14 (0.15)
	0.2	BH	23.42 (0.79)	23.43 (0.94)	23.45 (1.21)	23.54 (1.79)
		Bon	4.35 (0.07)	4.34 (0.11)	4.34 (0.15)	4.34 (0.20)
(10, 10, 10, 40)	0.1	BH	23.84 (1.18)	23.88 (1.34)	23.97 (1.64)	24.10 (2.42)
		Bon	2.32 (0.04)	2.32 (0.07)	2.32 (0.09)	(2.32 (0.12)
	0.2	BH	24.60 (0.81)	24.61 (0.93)	24.64 (1.19)	24.73 (1.78)
		Bon	4.68 (0.06)	4.68 (0.09)	4.68 (0.13)	4.67 (0.17)
(10, 10, 40, 40)	0.1	BH	24.53 (1.21)	24.58 (1.34)	24.68 (1.60)	24.82 (2.41)
		Bon	2.42 (0.03)	2.42 (0.05)	2.42 (0.08)	2.42 (0.10)
	0.2	BH	25.23 (0.81)	25.25 (0.94)	25.28 (1.18)	25.38 (1.79)
		Bon	4.86 (0.05)	4.86 (0.08)	4.86 (0.11)	4.86 (0.15)
(10, 40, 40, 40)	0.1	BH	23.84 (1.20)	23.88 (1.36)	23.96 (1.65)	24.08 (2.42)
		Bon	2.32 (0.04)	2.32 (0.06)	2.32 (0.09)	2.32 (0.12)
	0.2	BH	24.61 (0.82)	24.63 (0.95)	24.66 (1.20)	24.75 (1.80)
		Bon	4.68 (0.05)	4.68 (0.09)	4.68 (0.12)	4.68 (0.17)

confidence lower bounds (yellow longdashed line) are always the highest while the Bonferroni bounds (black solid line) are the lowest and the FCR-adjusted bounds (blue dotted line) in between.

7 Discussion

Since the groundbreaking paper of Hochberg and Benjamini (1995), the control of FDR is well studied when thousands or even millions genes are tested simultaneously. The statistical selective inference has been addressed in Benjamini and Yekutieli (2005) and is a hot topic now (see Benjamini, Heller, and Yekutieli (2009), Yekutieli (2012) and Benjamini and Bogomolov (2014) among others). Benjamini and Yekutieli suggested that it is important to account for selection when the selected intervals are reported or emphasized. We have applied Benjamini and Yekutieli (2005)'s procedure for constructing FCR-adjusted BH selected confidence intervals and the confidence lower bound for $\mu_k - \mu_1$ in Peng et al. (2008) to ensure FCR control and to take into account of the monotonicity of the dose-response studies in microarray experiments. The proposed interval estimation method provides more useful information than point estimates or test statistics about the selected genes' biological relevance and so is beneficial to understanding dose-response microarray experiments.

Acknowledgements The authors are grateful to the Editor and two anonymous referees for their comments and suggestions that led to substantial improvements in the manuscript. The first author's research was partially supported by NSERC of Canada.

Conflict of Interest

The authors have declared no conflict of interest.

Table 3 Simulated FCR of FCR-Adjusted BH and Bonferroni 95% CIs for BH and Bonferroni Level 0.05 Selection for Covariance with $\Sigma_{2,\rho}$.

Configuration	τ	Selection Method	ρ			
			0.0	0.2	0.5	0.9
(10, 20, 30, 40)	0.1	BH	0.00153	0.00177	0.001578	0.00163
		Bon	0.00009	0.00015	0.00015	0.00015
	0.2	BH	0.00353	0.00341	0.00358	0.00350
		Bon	0.00010	0.00015	0.00012	0.00010
(10, 10, 10, 40)	0.1	BH	0.00304	0.00294	0.00265	0.00286
		Bon	0.00019	0.00024	0.00020	0.00019
	0.2	BH	0.00635	0.00611	0.00609	0.00579
		Bon	0.00019	0.00025	0.00020	0.00022
(10, 10, 40, 40)	0.1	BH	0.00432	0.00417	0.00405	0.00427
		Bon	0.00025	0.00019	0.00019	0.00020
	0.2	BH	0.00882	0.00868	0.00844	0.00826
		Bon	0.00020	0.00022	0.00019	0.00019
(10, 40, 40, 40)	0.1	BH	0.00282	0.00296	0.00281	0.00276
		Bon	0.00014	0.00015	0.00015	0.00020
	0.2	BH	0.00655	0.00640	0.00629	0.00614
		Bon	0.00015	0.00017	0.00015	0.00015

References

Benjamini, Y. (2010) Simultaneous and selective inference: current successes and future challenges. *Biometrical Journal*, **52**, 708-721.

Benjamini, Y. and Bogomolov, M. (2014) Selective inference on multiple families of hypotheses. *J. R. Stat. Soc. Ser. B.* **76**, 297-318.

Benjamini, Y., Heller, R., and Yekutieli, D. (2009) Selective inference in complex research. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **1906**, 4255-4271.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B* **57**, 289-300.

Benjamini, Y., Krieger, A., and Yekutieli, D. (2006) Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* **93**, 491-507.

Benjamini, Y. and Liu, W. (1999) A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence. *Journal of Statistical Planning and Inference* **82**, 163-170.

Benjamini, Y. and Yekutieli, D. (2005) False discovery rate-adjusted multiple confidence intervals for selected parameters. *Journal of the American Statistical Association* **100**, 71-81.

Benjamini, Y. and Yekutieli, D. (2001) The control of the false discovery rate in multiple testing under dependency. *Ann. Statist* **29**, 1165-1188.

Bretz, F., Pinheiro, J. C., and Branson, M. (2005). Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics* **61**, 738-748.

Genz, A. and Bretz, F. (2009) *Computation of multivariate normal and t probabilities*. Lecture Notes in Statistics, Vol. 195. Springer, Heidelberg.

Hsu, J. C. and Berger, R. L. (1999). Stepwise confidence intervals without multiplicity adjustment for dose-response and toxicity studies. *Journal of the American Statistical Association* **94**, 468-482.

Hu, J., Kapoor, M., Zhang, W., Hamilton3, S. R., and Coombes1,K. R. (2005) Analysis of dose response effects on gene expression data with comparison of two microarray platforms, *Bioinformatics* **21**, 3524-3529.

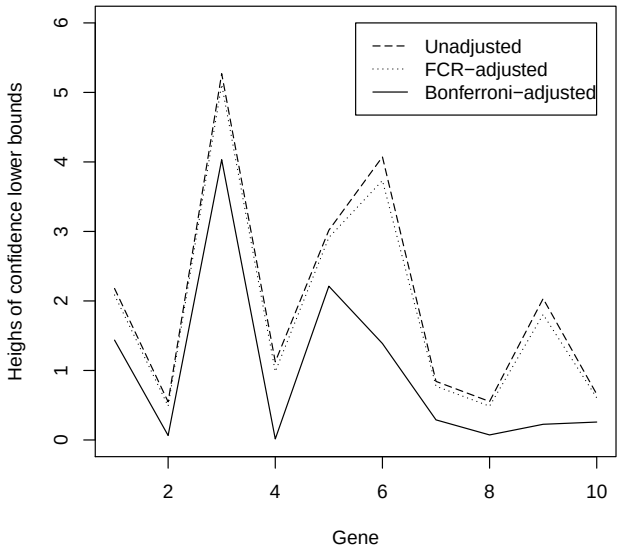

Figure 2 Heights of 10 confidence lower bounds for the first 10 DE genes selected by the Bonferroni procedure for $\mu_{g4} - \mu_{g1}$ under $\mu_{g1} \leq \cdots \leq \mu_{g4}$ using the unadjusted, FCR-adjusted, and Bonferroni approaches.

Jung, K., Friede, T., and Beibbarth, T. (2011) Reporting FDR analogous confidence intervals for the log fold change of differentially expressed genes. *BMC Bioinformatics* **12**, Article 288 (electronic).

Lin, Dan; Burzykowski, Tomasz; Perera, Tim; Shkedy, Ziv; Ghlmann, Hinrich W. H.; Geerts, Tamara; Yekutieli, Dani; De Bondt, An; Bijnens, Luc (2007). Testing for trends in dose-response microarray experments: a comparison of several testing procedures, multiplicity and resampling-based inference. *Stat. Appl. Genet. Mol. Biol* **6** Article 26 (electronic).

Lin, D.; Shkedy, Z.; Burzykowski, T.; Ion, R.; Ghlmann, H. W. H.; De Bondt, A.; Perera, T.; Geerts, T.; Van den Wyngaert, I.; Bijnens, L. (2008) An investigation on performance of significance analysis of microarray (SAM) for the comparisons of several treatments with one control in the presence of small-variance genes. Biom. J. **50**, 801-823.

Lin, D., Shkedy Z., Yekutieli, D., Amaratunga, D. and Bijnens, L. (2012) *Modeling Dose-Response Microarray Data in Early Drug Development Experiments Using R: Order-Restricted Analysis of Microarray Data.* Springer.

Miwa, T., Hayter, A., and Liu, W. (2000). Calculations of level probabilities for normal random variables with unequal variances with applications to Bartholomew's test in unbalanced one-way models. *Computational Statistics and Data Analysis* **34**, 17-32.

Peng, J., Lee, C. I. C., Davis, K., and Wang,W. (2008) Stepwise confidence intervals for monotone dose-response studies, *Biometrics* **64**, 877-885.

Robertson, T., Wright, F., and Dykstra, R. (1988). *Order Restricted Statistical Inference.* New York: Wiley.

Sarkar, S. K. (2002) Some results of the false discovery rate in stepwise multiple testing procedures. *Ann. Statist* **30**, 239-257.

Storey, J. D., Taylor, J. E. and Siegmund, D. (2004) Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *J. Roy. Statist. Soc. Ser. B* **66**, 187-205.

Tamhane, A. (2005) Comment for false discovery rate-adjusted multiple confidence intervals for selected parameters. *Journal of the American Statistical Association* **100**, 84-85.

Tamhane, A., Hochberg, Y., and Dunnett, C. W. (1996) Multiple test procedures for dose finding. *Biometrics* **52**, 21-37.

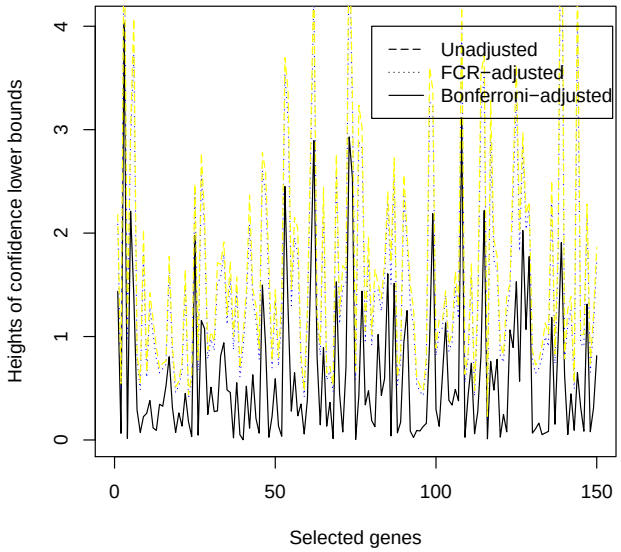

Figure 3 Heights of confidence lower bounds for the first 150 DE genes selected by the Bonferroni procedure for $\mu_{g4} - \mu_{g1}$ under $\mu_{g1} \leq \cdots \leq \mu_{g4}$ using the unadjusted, FCR-adjusted, and Bonferroni approaches.

Yekutieli, D. (2012) Adjusted Bayesian inference for selected parameters. *J. R. Stat. Soc. Ser. B.* **74**, 515-541.

Yekutieli, D. (2008) False discovery rate control for non-positively regression dependent test statistics. *J. Statist. Plann. Inference* **138**, 405-415.

